

Testentwicklung zur Erhebung der Schulterfunktion bei unfallchirurgischen und orthopädischen Patienten

**Von der Fakultät für Wirtschafts- und Sozialwissenschaften
der Universität Stuttgart zur Erlangung der Würde des Doktorgrades der
Philosophie (Dr. phil.) genehmigte Abhandlung**

vorgelegt von
Corinna Nerz (geb. Oberle)

aus Stuttgart

Hauptberichter: Prof. Dr. C. Becker

Mitberichter: Prof. Dr. W. Alt

Tag der mündlichen Prüfung: 15.02.2021

**Institut für Sport- und Bewegungswissenschaft
der Universität Stuttgart**

2020

Inhaltsverzeichnis

1	Zusammenfassung in deutscher Sprache	8
2	Abstract in englischer Sprache	11
3	Einleitung	13
4	Theoretische Aufbereitung des Problemfelds	17
4.1	Frakturepidemiologie	17
4.2	Frakturen der oberen Extremitäten.....	19
4.2.1	Anatomische Grundlagen und funktionelle Anatomie des Humerus	19
4.2.2	Inzidenz der Humerusfraktur.....	23
4.2.3	Diagnostik und Klassifikationen der Humerusfraktur.....	24
4.2.4	Risikofaktoren der Humerusfraktur	28
4.2.5	Behandlungskonzepte	29
5	Assessments der Schulterfunktionalität	33
5.1	Gütekriterien wissenschaftlicher Messinstrumente.....	33
5.2	Subjektive Messinstrumente.....	35
5.3	Subjektiv-objektive Messinstrumente	37
5.4	Objektive Messinstrumente	38
6	Fragestellungen	40
7	Testentwicklung und Evaluierung eines Tests zur Bestimmung der Schulterfunktion/ Publikationsübersicht	42
7.1	Entwicklung des modifizierten Wolf-Motor-Funktionstests (WMFT-O) für ältere Patienten nach Schulterverletzung und klinische Überprüfung der inter- und intra- rater Reliabilität (Manuskript 1).....	43
7.2	Anwendung des modifizierten Wolf-Motor-Funktionstests (WMFT-O) im Rahmen einer multizentrischen randomisierten, kontrollierten und teil-verblindeten Studie zur Generierung von Datensätzen zur Überprüfung der Testgütekriterien (Manuskript 2)	45

7.3	Überprüfung der Veränderungssensitivität und der inter-rater Reliabilität des videobasierten Wolf-Motor-Funktionstests (WMFT-O) sowie Korrelationsanalyse mit dem „Disability of the Arm, Shoulder and Hand Questionnaire“ (DASH) (Manuskript 3)	47
7.4	Ergebnisse einer multizentrischen randomisierten, kontrollierten und teilverblindeten Studie zur Überprüfung der Wirksamkeit eines additiven, assistiven Robotiktrainings nach proximaler Humerusfraktur (Manuskript 4)	49
8	Einordnung der Studienergebnisse in den Forschungszusammenhang	52
9	Schlussfolgerung und Ausblick	59
10	Literaturverzeichnis	62
11	Erklärung zum Eigenanteil der Dissertationsschrift	69
12	Weitere Publikationen und Kongressbeiträge	70
13	Danksagung	73
14	Anhang: Manuskripte 1-4	75

Abbildungsverzeichnis

Abbildung 1 Rechter Oberarmknochen (40)	20
Abbildung 2 Röntgenbild des rechten Schultergelenks (40)	22
Abbildung 3 Bewegungsmöglichkeiten im Schultergelenk und ausführende Muskeln (40)	23
Abbildung 4 AO- Klassifikation (45)	26
Abbildung 5 Klassifikation der Humeruskopffrakturen nach Neer (53)	27

Abkürzungsverzeichnis

AO-Klassifikation	A rbeitsgemeinschaft O steosynthese-Klassifikation
COPD	engl. „ C hronic o bstructive p ulmonary d isease“
CS	engl. „ C hange s core“
DASH.....	engl. „ D isability of A rm, S houlder and H and Questionnaire“
DGOU.....	D eutsche G esellschaft für O rthopädie und U nfallchirurgie
DGUV	D eutschen G esetzlichen U nfallversicherung
Engl.	in e nglischer Sprache
GOLD	engl. „ G lobal Initiative for Chronic O bstructive L ung D isease“
NYHA.....	engl. „ N ew Y ork H eart A ssociation“
PHF	p roximale H umerusfraktur(en)
r	Pearson Korrelationskoeffizient
RASTA.....	engl. „ R obotic A ssistive T raining A rmeo®Spring“
SES	engl. „ S tandardized e ffect s ize“
SOMC.....	engl. „ S hort O rientation- M emory- C oncentration Test“
SPADI.....	engl. „ S houlder P ain and D isability I ndex“
SRM.....	engl. „ S tandardized r esponse m ean“
VAS	V isuelle A nalog S kala (engl. „visual analogue scale“)
WMFT.....	W olf- M otor- F unktionstest
WMFT-O.....	W olf- M otor- F unktionstest- O rthopädie

I. Liste der wissenschaftlichen Veröffentlichungen zur publikationsbasierten Dissertation

Manuskript 1:

Oberle C, Becker C, Schölch S, Lenz JU, Studier-Fischer S, Augat P, Schwickert L. Inter-rater and intra-rater reliability of an adapted Wolf motor function test for older patients with shoulder injuries. *Z Gerontol Geriat.* 2017. DOI 10.1007/s00391-017-1226-2

Nachdruck mit Genehmigung (4867600636332) von "Copyright Clearance Center": Springer Nature; Zeitschrift für Gerontologie und Geriatrie; Inter-rater and intra-rater reliability of an adapted Wolf motor function test for older patients with shoulder injuries; Corinna Oberle et al. COPYRIGHT © 2017

Manuskript 2:

Nerz C, Schwickert L, Becker C, Studier-Fischer S, Müßig JA, Augat P. Effectiveness of robotic-assisted training added to conventional rehabilitation in patients with humeral fracture early after surgical treatment: study protocol of a randomised controlled trial. *Trials* 2017. DOI 10.1186/s13063-017-2274-z

This is an open access article. Permission to reuse this article is not required.

Manuskript 3:

Nerz C, Schwickert L, Schölch S, Gordt K, Nolte PC, Kröger I, Augat P, Becker C. Inter-rater reliability, sensitivity to change and responsiveness of the orthopaedic Wolf-Motor-Function-Test as functional capacity measure before and after rehabilitation in patients with proximal humeral fractures. *BMC Musculoskeletal Disorders* 2019. DOI 10.1186/s12891-019-2691-0

This is an open access article. Permission to reuse this article is not required.

Manuskript 4:

Kröger I, **Nerz C**, Schwickert L, Schölch S, Müßig JA, Studier-Fischer S, Nolte PC, Becker C, Augat P. Robot-assisted training after proximal humeral fracture: A randomised controlled multicentre intervention trial. *Clinical Rehabilitation.* 2021;35(2):242-252. DOI10.1177/0269215520961654 (*Original submission version*)

** Geteilte Erstautorenschaft*

Gratis reuse in dissertation or thesis, including where the dissertation or thesis will be posted in any electronic Institutional Repository or database.

Hinweis für die Leserin/den Leser:

Zugunsten eines besseren Leseflusses werden in der vorliegenden Arbeit im Regelfall Ausdrucksformen männlichen Geschlechts, wie Proband, verwendet. Es sei jedoch explizit darauf hingewiesen, dass dies selbstverständlich die weibliche Form, wie Probandin, miteinschließt.

1 Zusammenfassung in deutscher Sprache

Vor dem Hintergrund, dass die Anzahl der Frakturen der Schulter und des Oberarms stetig steigt, erhöht sich folglich auch der frakturbedingte Operations- und Rehabilitationsbedarf sowie als weitergehende Folgen die Pflegebedürftigkeit und die Mortalität. Die proximale Humerusfraktur (PHF) war 2017 mit rund 62.000 Fällen die am häufigsten in Deutschland registrierte Fraktur der oberen Extremität. Etwa 76% aller registrierten Fälle sind dabei dem Alterskollektiv der über 65-Jährigen zuzuschreiben. Gründe hierfür sind ein im Alter erhöhtes Sturzrisiko sowie gängige altersbedingte Knochenerkrankungen wie beispielsweise Osteoporose. Bei den rund 22% potentiell arbeitstätigen Patienten kann die PHF eine durch lange Immobilisationsphase und eine Vielzahl an Problemen, wie beispielsweise Muskelatrophie und Beweglichkeitsverlust der betroffenen Extremität, zu einer verminderten Arbeitsfähigkeit führen, welche häufig mit psychosoziale Probleme einhergeht. Die funktionelle Wiederherstellung der Schulterfunktion bleibt oft hinter den Erwartungen der betroffenen Patienten zurück. Folgen sind nicht nur soziale und wirtschaftliche Einschränkungen für die Patienten, sondern auch wirtschaftliche Belastungen für das Gesundheitssystem.

Die hier beschriebenen negativen Auswirkungen für die Patienten und die damit einhergehenden Belastungen für das Gesundheitssystem gilt es durch therapeutische Behandlungen bei früher Mobilisation in der funktionellen Nachbehandlung zu minimieren. Bislang fehlen jedoch standardisierte Leitlinien zur konservativen, operativen und therapeutischen Nachbehandlung von Patienten nach PHF. Um optimale Behandlungsschemata entwickeln zu können und die Rehabilitation von unfallchirurgischen und orthopädischen Patienten nach Schulterverletzungen im Allgemeinen zu standardisieren, muss die motorische Leistungsfähigkeit der Schulter zu unterschiedlichen Zeitpunkten objektiv beurteilt werden können. Dies erfordert neben patientenbasierten Fragebögen, auch objektive, klinischbasierte Messinstrumente. Schaut man jedoch in die Literatur, muss man feststellen, dass eben diese objektiven Messinstrumente bislang für unfallchirurgische und orthopädische Schulterpatienten fehlen.

Ein vielversprechender Ansatz für eine objektive Messung der Leistungsfähigkeit der oberen Extremität von neurologischen Patienten ist der Wolf-Motor-Funktionstest

(WMFT), welcher anhand unterschiedliche Bewegungsaufgaben, der zur Ausführung benötigten Zeit, der funktionellen Fähigkeit und der Bewegungsqualität die motorische Schulterfunktion beurteilt.

Manuskript 1 beschreibt die Modifikation der Testaufgaben und Bewertungskriterien des ursprünglichen WMFT und die damit einhergehende Anpassung auf unfallchirurgische und orthopädische Patienten und die Überprüfung hinsichtlich der Gütekriterien - Objektivität und Reliabilität. Dadurch entstand der klinische Wolf-Motor-Funktionstest-Orthopädie (WMFT-O). Dieser erweist sich als objektives und reliables Messinstrument mit sehr guter Inter- und Intra-Rater Reliabilität und weist zudem eine hohe interne Konsistenz auf. Demnach empfiehlt sich der WMFT-O zur Messung der Schulterfunktion in der Zielgruppe der älteren orthopädischen Patienten.

Aufbauend auf dem ersten Manuskript wird im 2. Manuskript eine multizentrische, randomisierte, kontrollierte und teil-verblindete Studie vorgestellt, bei welcher der WMFT-O als sekundärer Endpunkt zur Anwendung kommt. Hierbei sollen zudem Datensätzen zur Überprüfung der Testgütekriterien des WMFT-O in einer breiteren Zielgruppe generiert werden.

Manuskript 3 beschreibt die Überprüfung der Objektivität des videobasierten WMFT-O sowie die Analyse des Zusammenhangs der klinischen und der videobasierten Version des WMFT-O anhand der in Manuskript 2 gewonnen Daten. Zusätzlich wird die Veränderungssensitivität des WMFT-O und des bislang als Goldstandard gehandhabten subjektiven DASH-Fragebogens (engl. „Disability of Arm, Shoulder and Hand Questionnaire“) überprüft. Ergänzend wurde der Zusammenhang zwischen dem klinischen WMFT-O und dem DASH-Fragebogen analysiert. Der videobasierte WMFT-O erweist sich auch hierbei als objektives Messinstrument mit guter Inter-Rater-Reliabilität. Im Vergleich zur klinischen Version des WMFT-O ist die videobasierte Version hinsichtlich der Inter-Rater Reliabilität etwas niedriger. Die Veränderungssensitivität ist sowohl bei der klinischen Version des WMFT-O, als auch bei dem DASH-Fragebogen hoch. Jedoch korrelieren die Ergebnisse des klinischen WMFT-O nur schwach mit denen des DASH-Fragebogen.

In Manuskript 4 werden abschließend die Ergebnisse der in Manuskript 2 vorgestellten multizentrischen, randomisierten, kontrollierten und teil-verblindeten Studie präsentiert.

Insgesamt belegen die vorliegenden Studienergebnisse, dass der WMFT-O ein reliables und veränderungssensitives Instrument zur objektiven Messung der Leistungsfähigkeit bei unfallchirurgischen und orthopädischen Patienten nach PHF ist. Er bietet dadurch die Möglichkeit, objektiv den Operations- und Rehabilitationserfolg von unfallchirurgischen und orthopädischen Patienten zu messen und zu bewerten und schließt damit eine Forschungslücke in diesem Bereich. Weitere Studien sind notwendig, in welcher der WMFT-O in größeren randomisierten, kontrollierten Studien angewendet wird, um somit die Generalisierbarkeit der Ergebnisse zu erhöhen und gegebenenfalls anschließend eine Kurzversion des WMFT-O entwickeln zu können. Infolgedessen können dann längerfristig die Leitlinien zur Nachbehandlung nach PHF standardisiert und optimiert werden.

2 Abstract in englischer Sprache

In view of the fact that the number of fractures of the shoulder and upper arm is steadily increasing, the need for surgery and rehabilitation also increases, as well as the need for care and the mortality as further consequences. Proximal humeral fracture (PHF) was the most commonly registered upper limb fracture in Germany in 2017, with around 62,000 cases. About 76% of all registered cases are attributable to persons aged 65 years and older. This is due to an increased risk of falling as well as common age-related bone diseases such as osteoporosis. In the approximately 22% of potentially working patients, the PHF can lead to a reduced work ability due to a long immobilisation period and a variety of problems, such as muscle atrophy and loss of mobility of the affected arm, which is often associated with psychosocial problems. The functional recovery after a PHF is often disappointing. The consequences are not only social and economic constraints for patients, but also economic burdens on the healthcare system. Those negative effects for patients and the associated burdens on the healthcare system must be minimised by therapeutic treatments with early mobilisation and functional training.

To date, there is a lack of standardised guidelines for the conservative, surgical and therapeutic treatment of patients after PHF. In order to develop optimal treatment regimens and to standardise the rehabilitation of orthopaedic trauma patients after shoulder injuries in general, the motor function of the shoulder must be objectively assessed. In addition to patient-reported questionnaires, this also requires clinically based measuring instruments. However, in literature there is a lack of objective measuring instruments so far for orthopaedic trauma patients.

A promising approach for an objective assessment of the upper limb performance of neurological patients is the Wolf-Motor-Function-Test (WMFT), which assesses motor function of the shoulder by the required time, functional capacity and quality of movement of different movement tasks.

Manuscript 1 describes the modification and adaptation of the tasks and rating criteria of the original WMFT to orthopaedic trauma patients and the review of the quality criteria - objectivity and reliability. This resulted in the clinical Wolf-Motor-Functional-Test-Orthopaedics (WMFT-O). The WMFT-O proves to be an objective and reliable measuring instrument with very good inter- and intra-rater reliability. It has also a high

internal consistency. Accordingly, the WMFT-O is recommended as an assessment tool for measuring shoulder function in the target group of older orthopaedic trauma patients.

Based on the first manuscript, the second manuscript presents a multicentre randomised-controlled and partially-blinded study, in which the WMFT-O was used as a secondary outcome. In addition, datasets for testing the quality criteria of the WMFT-O in younger and older patients should be generated.

Manuscript 3 describes the review of the objectivity of the video-based WMFT-O, as well as the analysis of the relationship between the clinical and the video-based version of the WMFT-O based on the data obtained in manuscript 2. In addition, the sensitivity to change of the WMFT-O and the subjective DASH questionnaire (Disability of Arm, Shoulder and Hand Questionnaire), which was previously handled as a gold standard, is examined. Furthermore, the relationship between the clinical WMFT-O and the DASH questionnaire was analysed. The video-based WMFT-O proves to be an objective measuring instrument with good inter-rater reliability. Compared to the clinical version of the WMFT-O, the video-based version is somewhat lower in terms of inter-rater reliability. Sensitivity to change is high in both the clinical version of the WMFT-O and the DASH questionnaire. However, the results of the WMFT-O and the DASH questionnaire are weakly correlated.

In manuscript 4, the results of the multicentre randomised-controlled and partially-blinded study (manuscript 2) are presented.

Overall, the study results prove that the WMFT-O is a reliable and change-sensitive instrument for the objective measurement of motor performance in orthopaedic trauma patients after PHF. It offers the opportunity to objectively measure and evaluate the success of surgical and rehabilitation of orthopaedic trauma patients, thereby closing a research gap in this area. Further research is needed to apply the WMFT-O in larger randomized controlled trials, thus increasing the generalisability of the results and, when indicated, to develop a short version of the WMFT-O based on those new datasets. As a result, in longer term, the guidelines for post-treatment of PHF can be standardised and optimised.

3 Einleitung

Die Anzahl der Frakturen der Schulter und des Oberarms steigen stetig an. Wurden im Jahre 2000 noch rund 84.000 Schulter- und Oberarmfrakturen in deutschen Krankenhäusern behandelt, so waren es 2017 bereits 118.000 Fälle (1). Mit rund 62.000 registrierten Fällen in Deutschland war die proximale Humerusfraktur (PHF) 2017 die am häufigsten registrierte Fraktur der oberen Extremität (2). Auch ihre Inzidenz hat im letzten Jahrzehnt um mehr als 30% zugenommen. Waren es im Jahre 2000 in Deutschland noch 244 PHF pro 100.000 Einwohner, so ist diese Zahl bis zum Jahre 2017 auf 321 pro 100.000 Einwohner angestiegen (1,2). Dabei erleiden Frauen im Vergleich zu Männern fast dreimal so häufig eine PHF (2). In den Vereinigten Staaten ist eine ähnliche Inzidenz zu verzeichnen. Hier werden im Jahre 2030 rund 275.000 Fälle von PHF erwartet (3). Mit Hinblick auf die Altersverteilung bei PHF fällt auf, dass das Risiko eine PHF zu erleiden ab dem 45. Lebensjahr stark ansteigt und ab dem 65. Lebensjahr sich im Vergleich zu den Vorjahren noch einmal um das Vierfache erhöht (1). So wurden 2017 rund 76% aller PHF im Alterskollektiv der über 65-Jährigen verzeichnet. Laut Chu und Kollegen (2004) ist die PHF mit 10% die häufigste Verletzung des Oberarms in diesem Kollektiv (4). Gründe hierfür sind unter anderem die im Alter steigende Anzahl der Stürze (5), sowie die im Alter zunehmenden Häufigkeit der Osteoporose (6,7).

Die Konsequenzen aus der hier aufgezeigten, steigenden Inzidenz ist ein erhöhter Operations- und Rehabilitationsbedarf (8), wobei die funktionelle Wiederherstellung der Schulterfunktion oft hinter den Erwartungen zurückbleibt (9). Bedingt durch lange Immobilisationsphasen, sinkt die Muskelkraft und Beweglichkeit der betroffenen Extremität nach einer PHF rapide (10,11). Daher zählt die PHF auch zu den führenden Gründen für funktionelle Einschränkungen in den alltäglichen Aufgaben der betroffenen Patienten (12). Die hier beschriebene oftmals verzögerte Genesung und unvollständige Wiederherstellung der Schulterfunktion stellen nicht nur soziale und wirtschaftliche Einschränkungen für die Patienten, sondern auch wirtschaftliche Belastungen für das Gesundheitssystem dar (13,14). Bei den rund 22% potentiell arbeitstätigen Patienten¹ (1,2) kann die PHF durch eine Vielzahl an Problemen, wie

¹ Die Gruppe der potentiell arbeitstätigen Patienten bezieht sich auf das Merkmal Alter. Eingeschlossen wurden alle Personen zwischen 15 und 65 Jahren.

beispielsweise Muskelatrophie, Bewegungsverlust des Schultergelenks und der angrenzenden Gelenke, zu einer verminderten Arbeitsfähigkeit führen. Dies hat häufig auch psychosoziale Probleme zur Folge (9,10). Für die rund 76% der geriatrischen Patienten endet eine PHF mit einer schwerwiegenden Einschränkung der Selbständigkeit, was häufig eine Heimeinweisung nach sich bringt (15). Aufgrund des demografischen Wandels² wird sich das Problem der steigenden Institutionalisierung in den nächsten Jahren weiter verschärfen.

Die hier beschriebenen negativen Auswirkungen für den Patienten und die damit einhergehenden Belastungen für das Gesundheitssystem gilt es durch therapeutische Behandlungen bei früher Mobilisation in der funktionellen Nachbehandlung zu minimieren (16).

Sucht man jedoch in der Literatur nach Leitlinien und standardisierten Konzepten zur Nachbehandlung von PHF, so wird man nur in wenigen Fällen fündig. Bislang herrscht Uneinigkeit über die Dauer, Intensität und Frequenz von Therapien. Zudem beruhen die existierenden Leitlinien für die physio- und ergotherapeutische Nachbehandlung zumeist lediglich auf Expertenmeinungen (11). Ein aus der Neurologie stammender, neuer und vielversprechender Ansatz in der therapeutischen Nachbehandlung von Frakturen der oberen Extremität ist das assistive Robotiktraining (17). Es soll vor allem eine frühzeitige Mobilisation ermöglichen und somit die negativen Folgen der Immobilisation verringern (18–20).

Um den Rehabilitationserfolg im Allgemeinen und die Leistungsfähigkeit der betroffenen Extremität im Speziellen beurteilen zu können, werden bislang hauptsächlich nur subjektive Messinstrumente eingesetzt. Die deutsche Version (2.0) des DASH-Fragebogens (engl. „Disability of Arm, Shoulder and Hand Questionnaire“) erfasst die subjektive Einschätzung der Fähigkeit und die Beschwerden den betroffenen Arm im Alltag einzusetzen (21). Ein Assessment, das sowohl subjektive als auch objektive Elemente enthält, ist der Constant Score. Er beurteilt Schmerzen, Funktion und Kraft sowie das Bewegungsausmaß der betroffenen oberen Extremität

² Unter dem demografischen Wandel versteht man eine Veränderung der demografischen Bevölkerungsentwicklung hinsichtlich der Gesamtzahl an Mitgliedern und der zahlenmäßigen Struktur unter anderem bezogen auf die Altersgruppen und die Geschlechterverteilung. Bedingt durch den Geburtenrückgang bei gleichzeitigem Anstieg der Lebenserwartung steigt der Altersdurchschnitt der Bevölkerung weltweit an.

(22). Bislang gibt es kaum ein Messinstrument, welche die Leistungsfähigkeit des betroffenen Arms durch objektive Elemente beschreibt. Patientenbasierte Fragebögen haben den Nachteil, dass sie sowohl formale, als auch inhaltliche Antworttendenzen im Antwortverhalten, wie beispielsweise Angaben im Sinne der sozialen Erwünschtheit, abbilden. Die unter anderem dadurch entstehende Verzerrung in den Ergebnissen solcher subjektiver Messinstrumente zeigt die Notwendigkeit von objektiven, klinisch basierten Messinstrumenten auf (23,24). Diese sollen die Möglichkeit bieten, die funktionelle Leistungsfähigkeit des betroffenen Arms unabhängig vom Beobachter beziehungsweise vom Patienten zu beurteilen. Ein vielversprechendes objektives Messinstrument ist der Wolf-Motor-Funktionstest (WMFT) aus dem Jahre 1989. Er wurde von Dr. Steven L. Wolf und seinen Mitarbeitern von der Emory University School of Medicine in Atlanta in den Vereinigten Staaten entwickelt und umfasste in der ursprünglichen Version 21 Bewegungsaufgaben (engl. items) für die obere betroffene Extremität. In dieser Version wurde er für Schlaganfallpatienten oder Patienten, welche nach einer traumatisch bedingten Hirnverletzung unter einer Hemiparese leiden, entwickelt. Anhand des WMFT sollen die motorischen Fähigkeiten und Defizite dieser Patienten objektiv bewertet werden (25). Nach zwei durchgeführten Modifikationen wurde neben der benötigten Zeit zur Ausführung einer bestimmten Aufgabe auch die funktionelle Fähigkeit (26) und die Bewegungsqualität (27) mit in die Bewertung einbezogen. Um den WMFT auch bei stärker betroffenen Patienten einsetzen zu können, wurden zudem vier Aufgaben aus dem Test gestrichen, so dass die modifizierte Version von Bauder und Kollegen (2001) nur noch 17 Items umfasst (27).

Aufgrund der hier aufgezeigten Forschungslücke im Bereich der objektiven Messinstrumente zur Beurteilung der funktionellen Fähigkeit der betroffenen oberen Extremität bei unfallchirurgischen und orthopädischen Patienten, speziell nach PHF, bei gleichzeitig unabdingbarer Notwendigkeit solcher Messinstrumente, ist es das Ziel der vorliegenden Dissertation, den bereits existierenden WMFT, als objektives Maß für diese Patientengruppe zu modifizieren und zu evaluieren. Der Anteil der Bewegungsaufgaben, die vor allem mit dem proximalen Anteil des Armes ausgeführt werden, soll erhöht und gleichzeitig der Anteil der distalen, feinmotorischen Aufgaben reduziert werden, um so aus einem Messinstrument für ursprünglich neurologische Patienten, ein objektives, reliables und veränderungssensitives Messinstrument für unfallchirurgische und orthopädische Patienten zu entwickeln. Um diesen

Anforderungen gerecht zu werden und die durchgeführten Anpassungen evaluieren zu können, werden die Testgütekriterien des modifizierten Wolf-Motor-Funktionstests-Orthopädie (WMFT-O) überprüft. Im Anschluss daran werden die Ergebnisse einer randomisierten, kontrollierten Studie präsentiert, in welcher der WMFT-O zur Messung der funktionellen Wiederherstellung der Schulterfunktion nach PHF zum Einsatz kam.

Die vorliegende Dissertation ist dabei wie folgt gegliedert: Zunächst wird die Fraktur Epidemiologie im Allgemeinen und im Verlauf die Frakturen der oberen Extremität im Speziellen thematisiert. Unter diesem Punkt wird speziell auf die Anatomie, Inzidenz, Diagnostik, Klassifikation, Risikofaktoren und Behandlungskonzepte der Humerusfraktur eingegangen. Diese theoretische Aufbereitung des Problemfelds dient als Ausgangspunkt für die weiteren Überlegungen hinsichtlich spezieller Messinstrumente für die obere Extremität. Anschließend folgt ein Überblick über die bereits existierenden Assessments zur Beurteilung der Schulterfunktion, unterteilt in subjektive, subjektiv-objektive und objektive Messinstrumente, sowie eine Definition der Gütekriterien wissenschaftlicher Messinstrumente. Darauf folgend werden die Forschungsfragen formuliert, sowie die methodische Einbettung dieser Dissertation erläutert. Des Weiteren folgen die Zusammenfassungen der dissertationsrelevanten Manuskripte 1-4, welche sowohl die Modifizierung des WMFT, als auch die Beurteilung der Testgütekriterien (Objektivität, Reliabilität und Veränderungssensitivität), sowie die Ergebnisse der Anwendbarkeit des WMFT-Os in einer randomisierten, kontrollierten Studie beinhalten. Abschließend werden die Studienergebnisse in den Forschungszusammenhang eingeordnet und die Schlussfolgerungen aus der vorliegenden Dissertation gezogen. Zusätzlich wird ein Ausblick für zukünftige Forschungsansätze gegeben.

4 Theoretische Aufbereitung des Problemfelds

Im folgenden Abschnitt soll kurz auf die Frakturepidemiologie, die Frakturen der oberen Extremitäten im Allgemeinen und die Humerusfrakturen im Speziellen eingegangen werden.

4.1 Frakturepidemiologie

Frakturen sind bei älteren Personen meist die Folge von „Low-Energy-Traumen“. Hauptursachen sind eine reduzierte Knochendichte (Osteoporose) oder Stürze während alltäglicher Tätigkeiten (28). Bei jüngeren Patienten sind Hochenergietraumata, wie Sportverletzungen, Stürze bei sportlichen Aktivitäten oder Arbeitsunfälle, häufige Ursachen für eine Fraktur (29). Je nach Fallrichtung unterscheiden sich die aus einem Sturz resultierenden Frakturen. Hüftgelenksnahe Frakturen werden häufig durch Stürze auf die Seite verursacht, bei distalen Radius- und proximalen Humerusfrakturen (PHF) hat der Patient hingegen versucht, sich während des Sturzes mit seinem ausgestreckten Arm abzufangen (30,31). Das relative Risiko nach einer Fraktur eine weitere Fraktur zu erleiden ist bei Frauen fast doppelt und bei Männern mehr als doppelt so hoch (32). Die Femurfraktur (S72)³ ist mit 184.400 registrierten Fällen in Deutschland im Jahr 2016 die am häufigsten stationär behandelte Fraktur (1). Bei Personen über 65 Jahren tritt die Hälfte aller Femurfrakturen bei Menschen mit Pflegebedürftigkeit auf (33).

Betrachtet man die Hauptdiagnosen der Patienten, welche 2016 vollstationär in einem Krankenhaus in Deutschland aufgrund einer Fraktur aufgenommen wurden, so verteilen sich die Frakturinzidenzen wie folgt: ca. 184.000 Einlieferungen aufgrund einer Femurfraktur (S72), 133.000 aufgrund einer Unterarmfraktur (S52), 129.000 aufgrund einer Unterschenkelfraktur (S82), 116.000 aufgrund einer Oberarm- oder Schulterfraktur (S42), 100.000 aufgrund einer Becken- oder Lendenwirbelsäulenfraktur (S32), 71.000 aufgrund einer Rippenfraktur oder einer Fraktur im Bereich des Oberkörpers (S22), 41.000 aufgrund einer Schädel- oder Gesichtsfraktur (S02), jeweils 24.000 aufgrund einer Fraktur im Bereich des Fußes

³ „tenth revision of the International Classification of Diseases (ICD-10)“

(S92) und des Handgelenks (S62) und 10.000 aufgrund einer Fraktur der Halswirbelsäule (S12) (1).

Als Folge einer der oben genannten Frakturen starben 2018 in Deutschland insgesamt rund 10.000 Personen. Die Femurfraktur (S72) führt dabei mit 6.630 Fällen am häufigsten zum Tod. Als zweit häufigste Todesursache unter den Frakturen konnten die Becken- und Lendenwirbelsäulenfrakturen (S32) ausgemacht werden. Es folgten Frakturen im Bereich des Halses (S12), des Oberarms und der Schulter (S42), sowie des Schädels bzw. Frakturen im Bereich des Gesichts (34).

Die Studienergebnisse einer in Deutschland über vier Jahre durchgeführten Beobachtungsstudie, bei der die Studienpopulation fast 54.000 zuhause lebende Ältere über 65 Jahre umfasste, bestätigen die hier genannten Inzidenzen weitestgehend. Die Hüftfraktur (beinhaltet: die Schenkelhalsfraktur (S72.0), die pertrochantäre Fraktur (S72.1), sowie die subtrochantäre Fraktur (S72.2)) stellt auch in dieser Altersgruppe mit 36,6% der Studienpopulation die am häufigsten stationär behandelte Fraktur dar. Weitere Frakturinzidenzen verteilten sich wie folgt: 14,6% distale Radiusfrakturen (S52), 14,5% Frakturen der unteren Extremitäten (Fraktur des Femurschaftes (S72.3), distale Femurfraktur (S72.4), Tibia- und Fibulafraktur), 13,0% proximaler Humerusfrakturen (S42.2), 12,9% Wirbelsäulenfrakturen (S12), und 8,5% Beckenfrakturen (S32). Patienten mit Beckenfrakturen und Hüftfrakturen hatten das höchste Durchschnittsalter (82 Jahre), während Patienten mit Frakturen der unteren Extremitäten am jüngsten waren (76 Jahre) (12). Als Folge einer der oben genannten Frakturen starben im Alterskollektiv der über 65-Jährigen insgesamt 5.394 Patienten (10,0%) innerhalb der ersten sechs Monate nach der Fraktur und 5.507 Patienten (10,2%) wurden als Neuzugang in ein Pflegeheim aufgenommen. Verglichen mit einer Hüftfraktur, ist das Risiko nach einer Wirbelsäulenfraktur in ein Pflegeheim eingewiesen zu werden oder zu versterben am zweithöchsten. Das Mortalitäts- und Institutionalisierungsrisiko nach einer distalen Radiusfraktur ist hingegen im Vergleich am niedrigsten (12).

4.2 Frakturen der oberen Extremitäten

Frakturen der oberen Extremität (S42 und S52) sind insgesamt mit einem Anteil von etwas weniger als einem Drittel, nach den hüftgelenksnahen Frakturen (S72 und S32), die zweithäufigste Frakturgruppe (1). Dies ist auch im Alterskollektiv der über 65-Jährigen der Fall (12). Die frakturbedingten Auswirkungen auf das Gesundheitssystem bedingt durch eine eingeschränkte Mobilität und Selbständigkeit und ein damit einhergehendes erhöhtes Risiko der Institutionalisierung werden häufig unterschätzt (35). Im Vergleich zu den unteren Extremitäten, sind die Gelenke der oberen Extremitäten deutlich beweglicher. Dieser Unterschied beeinflusst die nach einem Trauma oder nach einer Fraktur angestrebte Zielsetzung der Therapie. Bei den unteren Extremitäten wird eine optimale Belastbarkeit und Stabilität angestrebt, wohingegen man bei den oberen Extremitäten auf eine volle Beweglichkeit und bei Handverletzungen im speziellen auf eine Wiederherstellung der sensiblen Feinmotorik abzielt (36).

Betrachtet man die Frakturen der oberen Extremität genauer, so lassen sich diese in zwei unterschiedliche Gruppen unterteilen: zum einen in die Frakturen im Bereich der Schulter und des Oberarms (S42.-) und zum anderen in die Frakturen im Bereich des Unterarms (S52.-). Zur ersten Gruppe gehören die Frakturen der Klavikula (S42.0), der Skapula (S42.1), des proximalen Endes des Humerus (S42.2), des Humerusschaftes (S42.3) sowie des distalen Endes des Humerus (S42.4). In der zweiten Gruppe können unter anderem Frakturen des proximalen Endes der Ulna (S52.0), des proximalen Endes des Radius (S52.1), des Ulnaschaftes (S52.2), des Radiuschafts (S52.3), eine kombinierte Ulna- und Radiuschaftfraktur (S52.4) und eine distale Fraktur des Radius (S52.5) zusammengefasst werden.

4.2.1 Anatomische Grundlagen und funktionelle Anatomie des Humerus

Anatomische Grundlagen

Der Humerus oder auch Oberarmknochen ist der größte und längste Knochen der oberen Extremität und einer der längsten Röhrenknochen des Menschen. An seinem oberen Ende befindet sich sein Kopf (Caput humeri). Er bildet zusammen mit der Gelenkspfanne (Glenoid) des Schulterblatts (Skapula) das proximale Gelenk der oberen Extremität – das Schultergelenk. Von Knorpel- und Bandstrukturen umschlossen wird es im Sinne eines Kugelgelenks zum beweglichsten Gelenk des

menschlichen Körpers. Distal ist der Humerus über das Ellenbogengelenk mit der Elle (Ulna) und Speiche (Radius) des Unterarms verbunden. Insgesamt kann der Humerus in vier Bereiche unterteilt werden. Von proximal beginnend unterscheidet man zwischen dem Kopf (Caput humeri), dem Hals (Collum humeri), dem Schaft (Corpus humeri) und dem distalen Endstück (Extremitas distalis). Der Humerusschaft ist der längste Teil des Humerus. Betrachtet man den proximalen Humerusabschnitt genauer, so kann dieser wiederum in vier Hauptfragmente unterteilt werden. Neben dem Humeruskopf, welcher schräg nach kranial, medial und dorsal gerichtet ist, liegt lateral davon der große (Tuberculum majus) und kleine Knochenhöcker (Tuberculum minus), die voneinander durch den Sulcus intertubercularis abgetrennt sind. Die Knochenhöcker bilden den Ansatzpunkt für die jeweiligen Sehnen der Rotatorenmanschette. Distal davon liegt der Hals des Humerus (37–39).

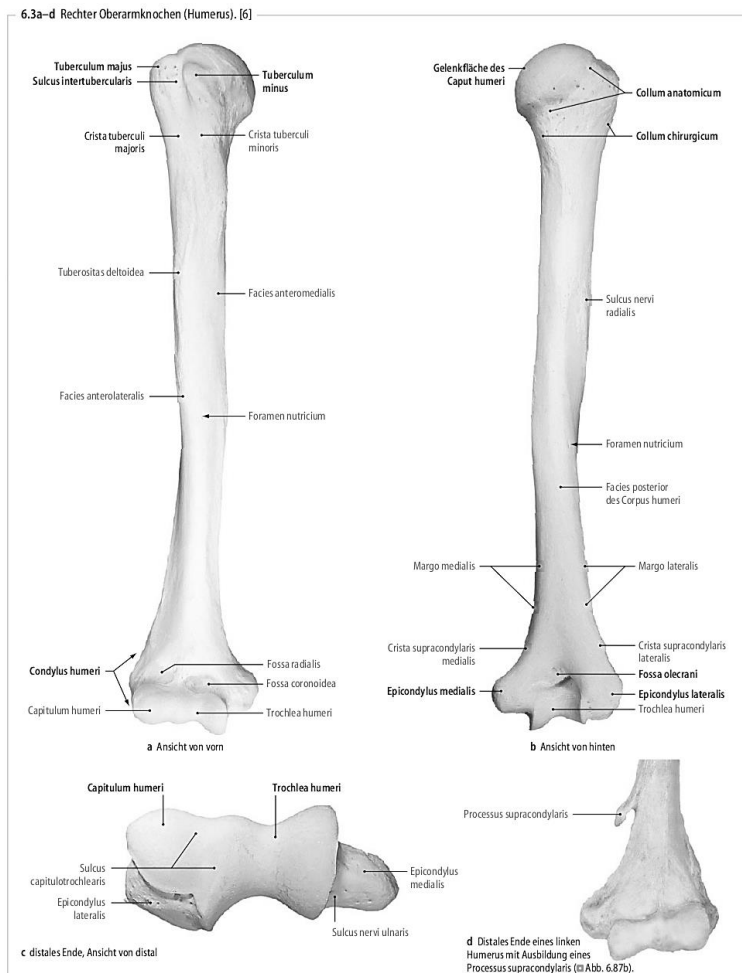


Abbildung 1 Rechter Oberarmknochen (40)

Nachdruck mit Genehmigung (4867600385326) von "Copyright Clearance Center": Springer Nature; Atlas der Anatomie des Menschen – Obere Extremität von Bernhard N. Tillmann COPYRIGHT © 2016

Funktionelle Anatomie

Das Schultergelenk ist eine funktionelle Einheit bestehend aus insgesamt fünf unterschiedlichen Gelenken. Hierzu zählen das Akromioklavikular-, das Sternoklavikular-, das Skapulothorakal- und das Glenohumeralgelenk, sowie die suprahumeralen bzw. subakromiale Fläche. Beim Skapulothorakalgelenk und der suprahumeralen Fläche spricht man jedoch auch von unechten Gelenken (39). Betrachtet man das Schultergelenk im Gesamten, so wird dies hauptsächlich durch das Zusammenspiel der beteiligten Muskeln bzw. deren Sehnen, die sogenannte Rotatorenmanschette, zusammengehalten. Durch die relativ flache Gelenkpfanne und dem Fehlen einer straffen Bandführung wird hierdurch ein hohes Maß an Beweglichkeit erreicht. Neben der stabilisierenden Aufgabe, sorgt die Schultermuskulatur auch für die Bewegungen im Schultergelenk. Sie verbindet den Humerus, die Knochen des Schultergürtels und den Thorax miteinander. Anatomisch unterscheidet man zwischen der skapulothorakalen, der thorakohumeralen und der skapulohumeralen Muskulatur. Bei der funktionellen Betrachtung des Schultergelenks muss ein besonderes Augenmerk auf die Rotatorenmanschette gelegt werden. Sie besteht aus vier Muskeln (M. subscapularis, M. supra- und infraspinatus und M. teres minor), welche am Humeruskopf ansetzen und diesen in der Schultergelenkpfanne zentrieren und stabilisieren. Mit Hilfe der Rotatorenmanschette wird der Oberarmkopf im Glenoid des Schulterblatts bewegt und somit die eigentliche Bewegung generiert. Betrachtet man die einzelnen Bewegungsrichtungen des Schultergelenks genauer, so lassen sich diese in folgende Gegensatzbeziehungen unterteilen (39–41):

1. Abduktion und Adduktion: Hierbei kommt es zu einer Bewegung in der Frontalebene um die Sagittalachse.
2. Anteversion und Retroversion: Hierbei kommt es zu einer Bewegung in der Sagittalebene um die Transversalachse.
3. Exorotation und Endorotation: Hierbei kommt es zu einer Bewegung in der Transversalebene um die Longitudinalachse.

6.10 Röntgenbild des rechten Schultergelenks einer 24-jährigen Frau im anteroposterioren Strahlengang. [10]

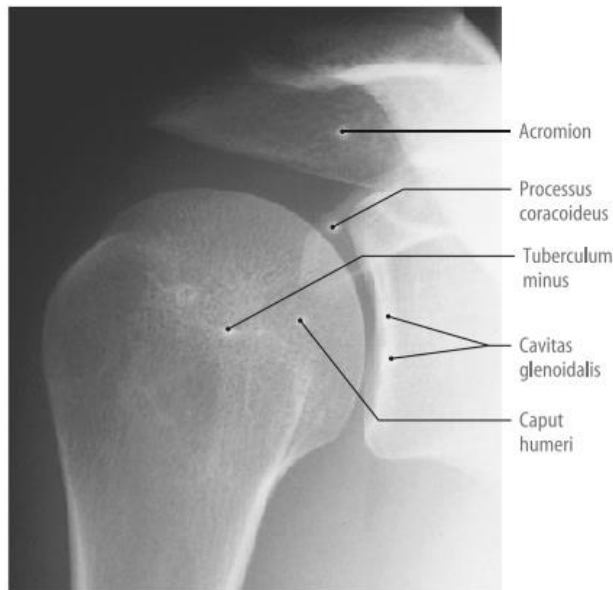


Abbildung 2 Röntgenbild des rechten Schultergelenks (40)

Nachdruck mit Genehmigung (4867600385326) von "Copyright Clearance Center": Springer Nature; Atlas der Anatomie des Menschen – Obere Extremität von Bernhard N. Tillmann COPYRIGHT © 2016

Hauptabduktoren sind der M. deltoideus und der M. supraspinatus. Der M. biceps brachii unterstützt jedoch die Bewegung. Bei einer adduzierenden Bewegung des Arms sind folgende Muskeln aktiv: M. pectoralis major, M. teres major und die Mn. rhomboidei, sowie M. latissimus dorsi und M. triceps brachii. Bei der Anteversion sind abhängig von der Phase unterschiedliche Muskeln beteiligt. Bei einer Anteversion bis zu 60° sind der M. deltoideus, M. coracobrachialis, M. biceps brachii und der M. pectoralis major beteiligt. Bei einer Anteversion zwischen 60-120° sind der M. trapezius und der M. serratus anterior beteiligt. Für eine einseitige Anteversion bis zu 180° sorgt hingegen der kontralaterale Anteil des M. erector spinae und eine beidseitige Anteversion wird über die Aktivierung beider Anteile des M. erector spinae bewirkt. Die vier wichtigsten Retroflexoren sind der M. latissimus dorsi, M. teres major und minor sowie der M. deltoideus. Für die Exorotation werden der M. teres minor und der M. infraspinatus aktiviert. Für die Endorotation wird hingegen der M. latissimus dorsi, M. teres major, M. pectoralis major und der M. subscapularis benötigt (39).

Das Bewegungsausmaß im Schultergelenk wird durch verschiedene Mechanismen limitiert. Bei einer Abduktion von 90° im Schultergelenk stößt der Humerus am Dach der Gelenkspfanne an. Wird jedoch eine Stellungsänderung der Schulterblattgelenkpfanne durch eine Drehung des Schulterblatts veranlasst, so kann dadurch der Arm bis zu 180° abduziert werden. Eine Adduktion ist bis zu 40° möglich. Der Arm kann bis zu 90° antevertiert und bis zu 40° retrovertiert werden. Kommt es in der Gelenkpfanne zu einer Stellungsänderung des Humeruskopfes, so kann die Anteversion in eine Elevation überführt werden. Hierdurch können Elevationswerte bis zu 170° erreicht werden. Exorotationen können bis zu 60° und Endorotationen bis zu 95° durchgeführt werden (6,40,41).

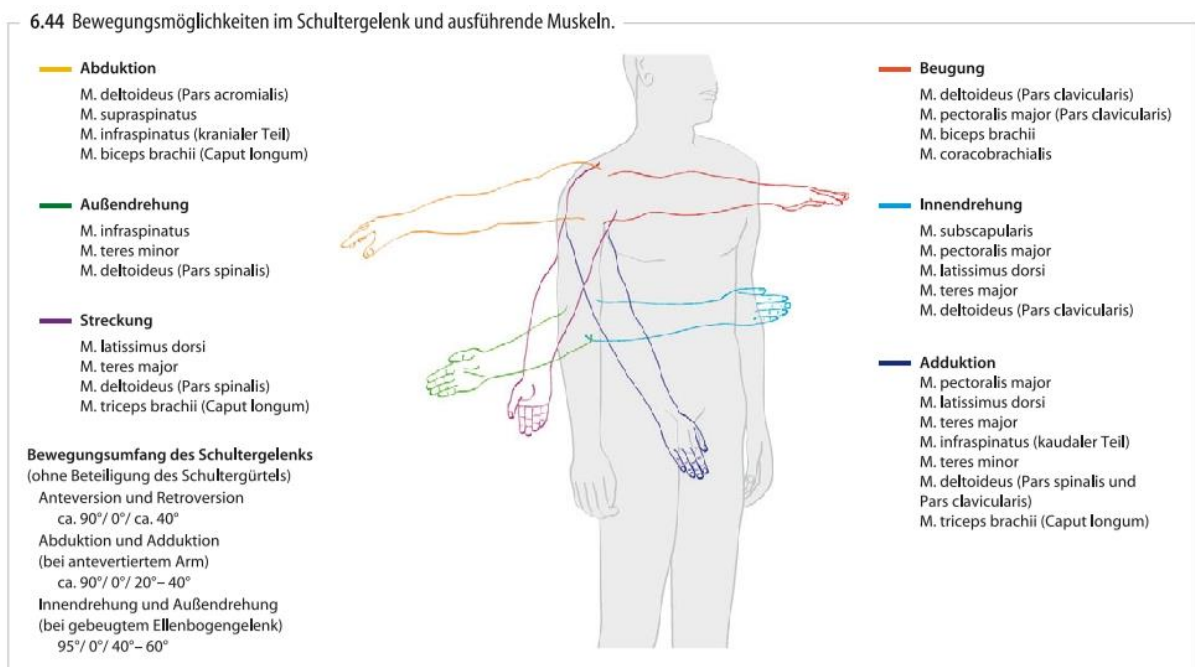


Abbildung 3 Bewegungsmöglichkeiten im Schultergelenk und ausführende Muskeln (40)

Nachdruck mit Genehmigung (4867600385326) von "Copyright Clearance Center": Springer Nature; Atlas der Anatomie des Menschen – Obere Extremität von Bernhard N. Tillmann COPYRIGHT © 2016

4.2.2 Inzidenz der Humerusfraktur

Im Jahre 2008 wurden in den Vereinigten Staaten rund 370.000 Krankenhausbesuche aufgrund von Humerusfrakturen verzeichnet. Hiervon betrafen rund 50% den proximalen Anteil des Humerus. Frauen waren dabei im Vergleich zu Männern doppelt so häufig betroffen. Bis zum Jahre 2030 soll die Zahl der Krankenhauseinweisungen in den USA aufgrund einer PHF auf 275.000 Fälle pro Jahr ansteigen (3). Bislang wurden für Europa wenige gesicherte epidemiologische Daten hinsichtlich Frakturen der oberen Extremitäten veröffentlicht (15). Aktuelle Studien aus Deutschland belegen

jedoch, dass 13% der frakturbedingten Krankenhauseinweisungen bei Personen über 65 Jahren aufgrund einer Humerusfraktur stattfinden (12). In der Altersgruppe der über 40-Jährigen betreffen 76% der Humerusfrakturen das proximale Ende (42). PHF und Handgelenksfrakturen gelten als die häufigsten Frakturen der oberen Extremität. Fasst man beide Frakturgruppen zusammen, so führen diese zu rund 30% der frakturbedingten Krankenhauseinweisungen (1). Seit 2000 ist die Inzidenz von PHF in Deutschland von 244 auf 321 pro 100.000 Einwohner pro Jahr angestiegen (1). Einer aktuellen deutschen Studie zur Folge liegt das Durchschnittsalter von Patienten, welche auf Grund einer PHF in ein Krankenhaus eingewiesen werden, bei 80,1 Jahren (32).

4.2.3 Diagnostik und Klassifikationen der Humerusfraktur

Um die Diagnose Humerusfraktur stellen zu können, sollte zunächst eine Unfallanamnese durchgeführt werden. Die Schonhaltung des betroffenen Arms eng am Oberkörper bzw. Brustkorb, Schwellungen und Hämatome, sowie Schmerzen im Allgemeinen oder speziell bei Druck auf den Humerus können vorab Anzeichen für eine Humerusfraktur sein (29). Im zweiten Schritt erfolgt eine detaillierte klinische Diagnostik. Hierbei wird die neurovaskulären Funktion überprüft (7). Dies umfasst die Überprüfung der Durchblutung, der Motorik und der Sensibilität. Des Weiteren ist es wichtig, die Funktion des Nervus axillaris zu untersuchen (29) und die Knochenqualität zu beurteilen (43). Zur Einschätzung der Frakturmorphologie und des Heilungsprozesses werden Röntgenbilder herangezogen. Hier ist es wichtig, dass mindestens zwei senkrecht aufeinander stehende Ebenen des proximalen Humerus abgebildet werden (30). Zum radiologischen Standard in der Diagnostik von Humerusfrakturen zählt idealerweise eine Traumaserie, welche eine true-a.p.-, Y-Aufnahme der Scapula und eine axiale Aufnahme umfasst. Die axiale Aufnahme ist jedoch in der Akutsituation oftmals schmerzbedingt nicht möglich (29). Die Computertomographie dient hauptsächlich zur präoperativen Diagnostik bei komplexen Frakturen zur Evaluation des geeigneten Operationsverfahrens (30). Sie gibt wichtige Zusatzinformation über die Anzahl, Größe und Lage der einzelnen Bruchstücke des proximalen Humerus (29).

Wurde bei einem Patienten eine Fraktur am Humerus diagnostiziert, so kann diese nach unterschiedlichen Kriterien beurteilt und klassifiziert werden. Die Klassifikation der Fraktur dient einerseits zur Dokumentation der Verletzung. So kann der

weiterbehandelnde Arzt nachvollziehen, ohne selbst eine eigenständige Diagnostik durchgeführt zu haben, um was für eine Fraktur es sich gehandelt hat. Andererseits können schnell Rückschlüsse auf die Verletzungsschwere gezogen werden. Neben der morphologischen Beschreibung der Fraktur ermöglicht die Klassifikation auch Hinweise auf die passende Therapie und die jeweilige Prognose (6).

Ein mögliches System der Klassifikation von Humerusfrakturen liefert die Arbeitsgemeinschaft für Osteosynthesefragen (AO-Foundation) – AO-Klassifikation oder auch Müller-Klassifikation (nach Maurice E. Müller benannt). Sie verwenden ein numerisches System, um die Beschaffenheit und die Lokalisation von Frakturen im Allgemeinen zu beschreiben. Die Humerusfraktur im Speziellen wird mit AO 1- deklariert, wobei die 1 für den gesamten Humerus steht. Die unterschiedliche Lokalisation am Humerus, an welcher sich die Fraktur befindet, wird jeweils mit einer weiteren Zahl benannt. Unter die Gruppe 1 fallen alle Frakturen am proximalen Ende de Humerus. Zur Gruppe 2 gehören alle Frakturen, welche den Knochenschaft (Diaphysis⁴) betreffen und zur Gruppe 3 zählen alle Frakturen, welche sich am distalen Ende des Humerus befinden. Durch weitere Untergruppen (Gruppe A, B und C), kann der Frakturort im jeweiligen Knochensegment noch genauer beschrieben werden. Der Grad der Segmentlokalisierung, die Segmentzahl und die Dislokationsrichtung werden wiederum durch weitere Untergruppen (1, 2 und 3) beschrieben (44).

⁴ Diaphysis: Von Metaphyse und Epiphyse abzugrenzendes Mittelstück der Röhrenknochen, das auch als Schaft bezeichnet wird.

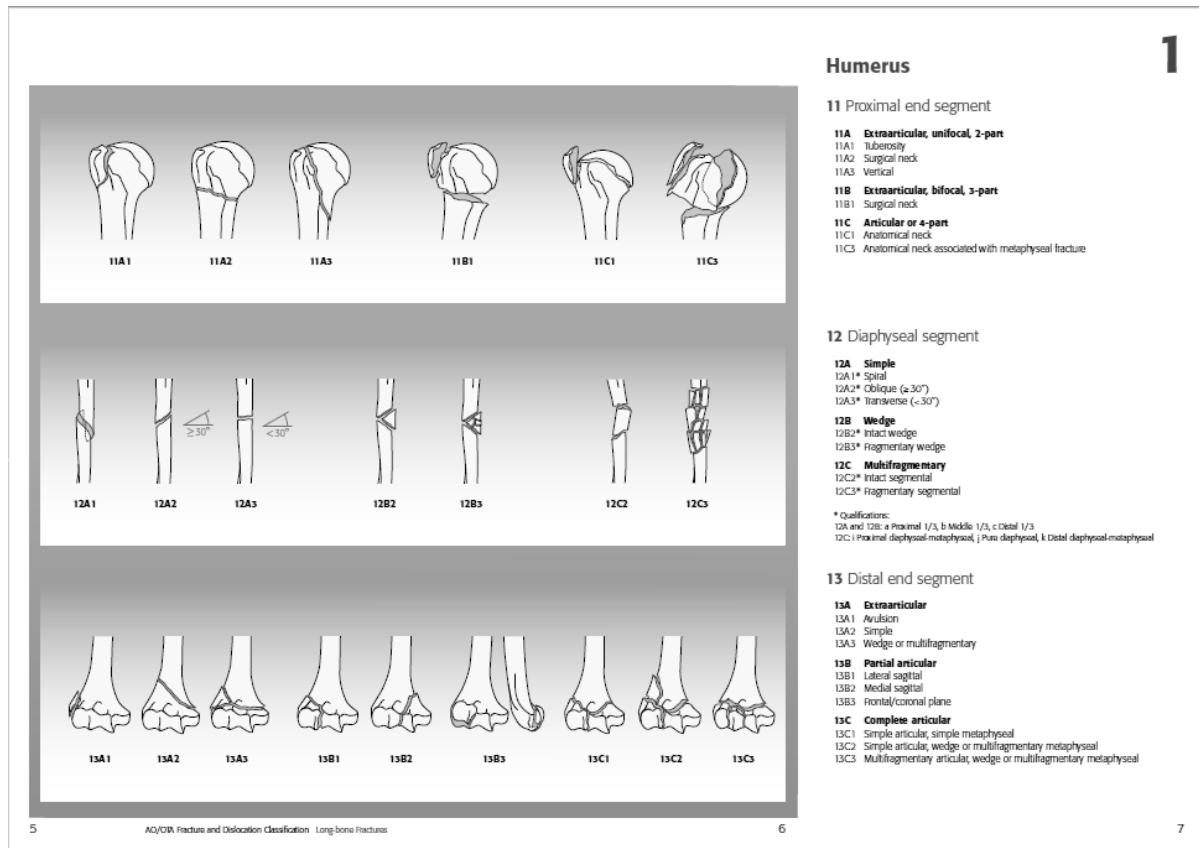


Abbildung 4 AO- Klassifikation (45)

"Copyright © 2018 by AO Foundation, Switzerland"

Laut Nellißen ist die AO-Klassifikation nicht ausreichend reproduzierbar (5) und hat im klinischen Alltag nur einen geringen Stellenwert. Am gebräuchlichsten ist die Klassifikation von Neer (46,47). Diese baut auf Codeman (1934) auf und hat sich im klinischen Alltag weitgehend durchgesetzt (31,48). Neer übernimmt die Einteilung von Codeman und erweitert die Klassifikation hinsichtlich des Dislokationsgrads und der Vaskularität (5,46,49,50). Die Klassifikation geht davon aus, wie viele Fragmente vorliegen. Verläuft die Frakturlinie am Collum anatomicum, so spricht man von einer Fraktur im 1. Segment. Bei einer Fraktur im 2. Segment verläuft die Bruchkante am Tuberculum minus, beim 3. Segment im Tuberculum majus und bei einer Fraktur im 4. Segment verläuft die Frakturlinie im Collum chirurgicum (5). Neer bezeichnet eine Fraktur als nicht-disloziert oder aber auch als 1-Segmentfraktur, wenn entweder die Verschiebung der Segmente kleiner als 1 cm sind oder aber die Winkelschiefstellung (Angulation) die 45° Grenze nicht überschreitet. Werden diese Grenzwerte überschritten, so liegt nach Neer eine dislozierte Fraktur vor. Von einer 2-Segmentfraktur spricht Neer, wenn ein Segment disloziert ist, alle weiteren Segmente jedoch physiologisch korrekt zueinanderstehen. Äquivalent hierzu sind bei einer 3-

Segmentfraktur zwei Segmente disloziert. Wenn alle vier Segmente disloziert sind, spricht man von einer 4-Segmentfraktur (5).

Lill und Josten (2001), sowie Tingart und Kollegen (2001) empfehlen eine kombinierte Betrachtung der Klassifikationen zur Bestimmung und Klassifikation von Frakturen (48,51).

Wirft man einmal einen Blick auf die Verteilung der PHF, so fällt auf, dass die 1-Segmentfrakturen mit etwa 80% überwiegen. Bei nur etwa 10% aller Fälle liegt eine 2-Segmentfraktur vor und in jeweils etwa 4% der Fälle eine 3- bzw. 4-Segmentfraktur (52). Zusammenfassend lässt sich sagen, dass eine schlechtere Behandlungsprognose gestellt werden muss, sobald jedes einzelne Segment um mehr als 10 mm disloziert ist und eine Angulation um mehr als 45° und somit eine höhere Frakturkategorie vorliegt (30). Des Weiteren fällt mit steigender Anzahl dislozierter Segmente die Prognose für den Patienten ungünstiger aus.
















I 	DISLOZIERTE FRAKTUREN		
	Anzahl der Fragmente		
	2	3	4
Minimale Dislokation			
II Collum anatomicum 			
III Collum chirurgicum 			
IV Tuberculum majus 			
V Tuberculum minus, 			
VI anteriore Luxationsfraktur 			
posteriore Luxationsfraktur 			

Abbildung 5 Klassifikation der Humeruskopffrakturen nach Neer (53)

Nachdruck mit Genehmigung (4867600068994) von "Copyright Clearance Center": Springer Nature; Der Unfallchirurg; Die Humeruskopffraktur; P Habermeyer COPYRIGHT © 1997

4.2.4 Risikofaktoren der Humerusfraktur

Als Hauptrisikofaktor für Humerusfrakturen im Allgemeinen und proximale Humerusfrakturen (PHF) im Speziellen, ist eine reduzierte Knochendichte im Rahmen einer osteoporosen Erkrankung zu nennen. Osteoporose zählt laut einer Definition der Weltgesundheitsorganisation zu den systemischen Erkrankungen des Skeletts (28). Hierbei reduziert sich mit zunehmendem Alter die Knochendichte und eine Verschlechterung der Mikroarchitektur des Knochens tritt auf (7). Folgen sind eine erhöhte Knochenfragilität und ein erhöhtes Frakturrisiko. Insgesamt leiden 7,8 Millionen Menschen in Deutschland unter Osteoporose, darunter 6,5 Millionen Frauen. Die Prävalenz nimmt ab dem Beginn der sechsten Lebensdekade deutlich zu. Jede zweite Frau und jeder fünfte Mann erleidet im Laufe des Lebens eine osteoporotisch bedingte Fraktur (28). Die erhöhte Prävalenz der Osteoporose bei Frauen erklärt unter anderem auch die weibliche Dominanz bei PHF (7).

Als ein weiterer, primärer Risikofaktor für eine Humerusfraktur ist das Sturzrisiko zu nennen (5). Gangunsicherheiten und daraus resultierende Stürze sind vor allem in der Geriatrie ein Problem mit hohem Stellenwert (15). Bei jüngeren Personen treten Stürze auch häufig bei Sport- und Freizeitaktivitäten oder im Arbeitsumfeld auf. Nach PROFANE (Prevention of Falls Network Europe) ist „ein Sturz [...] ein unerwartetes Ereignis, wobei der Betroffene auf dem Boden oder auf einer tieferen Ebene zum Stillstand kommt (54).“ Betrachtet man ältere Menschen über 65 Jahre, so stürzen 30% von ihnen mindestens einmal pro Jahr und wiederum 30-40% von ihnen erleiden schwerwiegende Verletzungen (28). Des Weiteren nimmt die Häufigkeit der Stürze mit jeder Lebensdekade um etwa 10% zu (15). Dabei liegt die Anzahl der Stürze bei Frauen allgemein höher als die bei Männern (55). Was wiederum ein weiterer Grund für die weibliche Dominanz bei PHF ist. Die Mechanismen, welche bei einem Sturz zu einer PHF führen, sind meist Stürze auf die aus Eigenschutz ausgestreckte Hand, auf die Schulter bei angelegtem Ober- und Unterarm oder ein Direkttrauma (30,31). Ein Sturz bei einem außenrotierten, abduzierten Arm führt bei jüngeren Patienten häufig noch, aufgrund der im Vergleich zum Kapsel-Band-Apparat stärkeren Knochenstruktur, zur einer Schultergelenkluxation, wohingegen ein solcher Sturz im höheren Alter meistens mit einer PHF endet (30).

Weitere sekundäre Risikofaktoren sind unter anderem chronischer Alkoholabusus, eine Krankheit, welche ähnlich der Osteoporose zur Abnahme der Knochen-

mineraldichte und Knochenmasse führt, Hypogonadismus, eine endokrine Funktionsstörung der Hoden, bei welcher es zu Testosteronmangel kommt, Langzeit-Kortikoid-Therapien und verschiedene gastrointestinale Erkrankungen (7). Später wurden dieser Liste noch Arthrose, Degeneration der Rotatorenmanschette und Multimorbidität hinzugefügt (28).

4.2.5 Behandlungskonzepte

Bislang fehlen evidenzbasierte Leitlinien und Therapieschemata zur Nachbehandlung von Humerusfrakturen. Die Studienlage der kontrollierten randomisierten Studien zu diesem Thema ist unzureichend (11). In den wenigen existierenden Studien zeigt sich ein Trend hin zur konservativen Therapie, speziell bei älteren Menschen (56). Aufgrund von geringen Fallzahlen kann daraus jedoch noch keine Empfehlung ausgesprochen werden. Bei jüngeren Patienten sind die anatomische Repositionierung und die osteosynthetische Stabilisierung anzustreben, wohingegen bei älteren Patienten auch die endoprothetische Versorgung berücksichtigt und diskutiert werden sollte. Zudem müssen bei älteren Menschen altersbedingte Besonderheiten, wie beispielsweise generalisierte Arteriosklerose (31) oder Zusatzverletzungen (30) bzw. Begleiterkrankungen berücksichtigt werden (29). Denn in Folge operativ indizierter Störungen können Patienten ihre Eigenständigkeit verlieren, was wiederum zur permanenten Betreuung in einer Pflegeeinrichtung oder sogar zum Tod führen kann (16). Dennoch bleibt, aufgrund der aktuellen Datenlage, die Behandlung von Humerusfrakturen individuell und muss vor Therapiebeginn mit dem Patienten, in Hinblick auf seine individuellen Bedürfnisse, abgesprochen werden. Abhängig von der gewählten Behandlungsstrategie kann es unter anderem zu Komplikationen wie Schultersteifigkeit, Humeruskopfnekrose, Infektionen, Bewegungseinschränkungen und Schmerzen kommen (29).

4.2.5.1 Konservative Behandlung

Die konservative Behandlung ist bei nicht-dislozierten Frakturen meist ausreichend. Es erfolgt eine Ruhigstellung durch einen Schulterverband (31). Bei dislozierten Frakturen hingegen wird es schwieriger eine klare Empfehlung für oder gegen eine konservative Behandlung auszusprechen. Die Entscheidung muss hier individuell getroffen werden. Allgemein kann man jedoch sagen, dass man sich bei älteren Patienten bezüglich der Operationsindikation eher an den Grenzwerten von Neer (<1

cm, $<45^\circ$; (46,50)), bei jüngeren Patienten eher die Grenzwerte von Lill ($<0,5$ cm, $<20^\circ$, $<0,2$ cm Tuberkuladislokation) orientiert (48). Die 60-80% nicht oder nur gering dislozierten Frakturen werden, vor dem Hintergrund des Fortschritts in der Osteosynthese, hingegen den Empfehlungen von Neer und Lill größtenteils operativ versorgt (29). Die zu erwartenden Ergebnisse einer konservativen Behandlung bei nicht-dislozierten Frakturen sind gut. Die Schulterbeweglichkeit kann meist bis zu 85% der Gegenseite erreichen. Mögliche Komplikationen einer konservativ behandelten PHF sind Schmerzen, Einschränkungen im Bewegungsausmaß und Humeruskopfnekrosen⁵ (29).

4.2.5.2 Operative Behandlung

Aufgrund des Fortschritts im Bereich der operativen Versorgung mit winkelstabilen Implantaten, ist in den letzten Jahren ein eindeutiger Trend weg von der konservativen, hin zur operativen Versorgung der Humerusfraktur zu beobachten. Instabile oder stark dislozierte Frakturen, welche die Grenzwerte von Neer (<1 cm, $<45^\circ$) und Lill ($<0,5$ cm, $<20^\circ$, $<0,2$ cm Tuberkuladislokation) überschreiten, sollten operativ versorgt werden. Weitere Operationsindikationen sind Nerven- oder Gefäßverletzungen, offene oder Luxationsfrakturen, sowie Head-split- oder Collum-anatomicum-Frakturen (29). Allgemein kann man zwischen offenen oder geschlossenen Repositionen bzw. nicht-minimalinvasiven und eingeschobenen (minimal-invasiven) Operationsverfahren unterscheiden. Burkhart und Kollegen (2013) unterscheiden bei der operativen Behandlung der PHF prinzipiell zwischen kopferhaltenden und kopfersetzenden operativen Versorgungsarten (29). Unter die kopferhaltenden Techniken fällt die osteosynthetische Versorgung. Sie stellt heutzutage das am häufigsten durchgeführte operative Behandlungsverfahren der PHF dar (29). Hierbei werden zwei oder mehrerer Knochen bzw. Knochenfragmenten mit dem Ziel der anatomischen Reposition und der stabilen Fixation der Fraktur operativ miteinander verbunden. Man kann zwischen Draht-, Schrauben-, Marknagel-, Verriegelungsnagel- und Plattenosteosynthesen unterscheiden. Die operative Versorgung mit winkelstabilen Osteosynthesen verlaufen bei jüngeren Patienten weitestgehend komplikationslos, wohingegen bei älteren Menschen mit osteoporotischen Frakturen etwas häufiger Komplikationen auftreten. Oftmals ist die Versorgung durch eine winkelstabile Plattenosteosynthese zu rigide

⁵ Humeruskopfnekrose: Absterben von Gewebezellen durch eine Minderdurchblutung am Humeruskopf

und es kommt zum sogenannten „Cutting Out“, bei welchem es durch die winkelstabile Position der Schraube zu einem Herausbrechen eines Knochenfragments aus dem Humerus kommt. Das Osteosyntheseversagen ist jedoch auch durch eine konsekutive Humeruskopfnekrose bedingt. Die Komplikationsrate der winkelstabilen Plattenosteosynthesen liegt bei etwa 25%. Die Endoprothese, ein Implantat, welches das geschädigte Gelenk teilweise oder ganz ersetzt, zählt zu den kopfersetzenden Techniken. Operationsindikationen für eine Endoprothese sind unter anderem eine fortgeschrittene Osteoporose, eine zuvor fehlgeschlagene osteosynthetische Versorgung oder die dislozierte Mehrfragmentfraktur insbesondere des älteren Menschen. Ist die Rotatorenmanschette bereits vor der Fraktur beschädigt, wie es bei älteren Menschen häufig der Fall ist, so zeigen sich nach Implantation einer Frakturoendoprothese oft schlechte Ergebnisse hinsichtlich der Schulterfunktion. Inverse Prothesen sind in diesen Fällen unbedingt zu verwenden. Dennoch sind in der Literatur Hinweise darauf zu finden, dass nach endoprothetischer Frakturversorgung Einschränkungen in der Schulterfunktion und dem Bewegungsumfang zurückbleiben (29).

4.2.5.3 Therapeutische Nachbehandlungskonzepte

Auch die Nachbehandlungskonzepte unterscheiden sich bezogen auf die vorangegangene Behandlungsmethode. Zudem gibt es häufig von Klinik zu Klinik individuelle Unterschiede bezüglich der Details der Nachbehandlung. Allgemein lässt sich jedoch sagen, dass bei einer konservativen Behandlung einer Humerusfraktur in der 1. Woche nach der Fraktur eine Ruhigstellung durch einen Gilchrist-Verband empfohlen wird. Es kann bereits damit begonnen werden, die angrenzenden Gelenke (Ellenbogen- und Handgelenk) zu beüben. Zudem wird eine Ödemprophylaxe durch beispielsweise manuelle Lymphdrainage empfohlen. In der 2. und 3. Woche kann mit Pendelbewegungen mit vorgeneigtem Oberkörper begonnen werden. Das Tragen des Gilchrist-Verbands sollte sich nur noch auf die Nacht beschränken. Von der 4. bis einschließlich der 6. Woche kann mit aktiv-assistiv Bewegungen innerhalb der physiotherapeutischen Behandlung begonnen werden. Diese sollten eine Abduktion/Anteversion von 90° jedoch noch nicht überschreiten. Die Muskelkräftigung sollte ab hier im Fokus der Therapie stehen. Ab der 7. Woche kann die Schulter für aktive Bewegungsübungen freigegeben werden. Die Belastung auf den Humerus sollte trotz Bewegungsfreigabe langsam und schrittweise gesteigert werden (29). Der

Deutschen Gesellschaft für Orthopädie und Unfallchirurgie (DGOU) zufolge ist eine konservativ behandelte proximale Humerusfraktur ab der 6. Woche belastungs- und ab der 12. Woche trainingsstabil (57). Zwischen den einzelnen Behandlungsschritten sind Röntgenaufnahmen obligat.

Die operativ versorgte PHF ist ab dem Operationstag bewegungs-, ab der 6. Woche belastungs- und ab der 12. Woche trainingsstabil. Die DGOU empfiehlt ab dem 1. postoperativen Tag assistive Bewegungen distal der Schulter sowie manuelle Lymphdrainage. Bis zum 3. postoperativen Tag sollte eine Kontrolle der Operationsergebnisse anhand einer Röntgenaufnahme geschehen. Bis zur 2. postoperativen Woche soll assistives/aktives Bewegen in alle Richtungen mit in die Therapie eingebaut werden. Für die Flexion und die Abduktion besteht eine Limitation von 90° und die Rotation soll nur im schmerzfreien Bereich stattfinden. Bis zur 6. Woche soll das aktive Bewegen in alle Richtungen ohne Limitation Inhalt jeder physiotherapeutischen Behandlung sein. Ab der 8. Woche kann intensives Muskelaufbautraining unter dem Einsatz von Geräten und bis zur 12. Woche Funktionstraining hinzugefügt werden. Ab 3. Monaten ist ein sportartspezifisches Training wieder möglich (57).

5 Assessments der Schulterfunktionalität

Bereits 1987 wollten die Mediziner und Autoren Constant und Murley eine klare Trennung zwischen der diagnostischen und der funktionellen Bewertung der Schulter schaffen. Bezieht sich eine diagnostische Bewertung meist auf die Anamneseerhebung, die körperliche Untersuchung, sowie bildgebende Verfahren (Röntgenbild, Computer- oder Kernspintomographie), so können daraus nicht immer ausreichend Erkenntnisse bezüglich des Funktionsgrades der Schulter gewonnen werden. Daher soll zur Beurteilung des funktionellen Fortschritts nach Verletzungen auch die Bewertung der Schulterfunktion herangezogen werden (22). Im englischsprachigen Raum spricht man auch von „clinical outcome assessments“. Sie dienen dazu einen Behandlungsnutzen direkt oder indirekt zu messen (58). Um die Funktionalität und die Kapazität der Schulter, sowie die Aktivität beurteilen zu können, werden sowohl patientenbasierte Fragebögen, als auch objektive, klinischbasierte Messinstrumente benötigt (23,24).

Im folgenden Abschnitt soll kurz auf die Gütekriterien wissenschaftlicher Messinstrumente eingegangen werden und die zurzeit verfügbaren und genutzten Assessments zur Erhebung und Messung der Schulterfunktionalität vorgestellt werden.

5.1 Gütekriterien wissenschaftlicher Messinstrumente

Gütekriterien dienen der Qualitätsbeurteilung wissenschaftlicher Daten, weitgehend der Erhebung und Analyse quantitativer Messergebnisse und der Qualitätssicherung in der Forschung. Jedes wissenschaftliche Messinstrument muss den Gütekriterien innerhalb eines gewissen Referenzbereiches genügen. Man kann hierbei zwischen Haupt- und Nebengütekriterien unterscheiden. Zu den drei allgemeingültigen Hauptkriterien zählen die Objektivität, die Reliabilität und die Validität eines Messinstruments (59).

Die **Objektivität** gibt an, inwieweit das Ergebnis eines Tests unabhängig von ungewollten Einflüssen durch die erhebende Person zustande gekommen ist. Ein Maß für die Objektivität eines Messinstruments ist die Inter-rater Reliabilität. Sie überprüft die Übereinstimmung der Ergebnisse von unterschiedlichen erhebenden Personen.

Wenn mehrere Untersucher (Rater) mit einem Messinstrument bei dem gleichen zu untersuchenden Sachverhalt zum exakt gleichen Ergebnis kommen, so ist das Messinstrument im höchsten Maß objektiv (59).

Die **Reliabilität** oder auch Zuverlässigkeit gibt den Grad der Genauigkeit an, mit der ein Messinstrumentes einen bestimmten Sachverhalt misst, unabhängig davon, ob das Messinstrument diesen Sachverhalt auch zu messen beansprucht. Die Reliabilität eines Messinstruments ist somit ein Indikator für die Replizierbarkeit der gewonnenen Ergebnisse. Ein Maß zur Bestimmung der Reliabilität eines Messinstruments ist die Intra-rater Reliabilität. Hierbei wird ein Messinstrument zu zwei unterschiedlichen Messzeitpunkten von demselben Untersucher (Rater) unter gleichen Bedingungen und bei demselben Sachverhalt/Untersuchungsgruppe wiederholt angewendet (59). Dies geschieht meist über ein Test-Re-Test Design.

Die **Validität** oder auch Gültigkeit gibt an, inwieweit ein Messinstrument das Merkmal beziehungsweise den Sachverhalt misst, dass es zu messen beansprucht. Ein Maß für die Validität eines Messverfahrens sind Vergleiche (Korrelationsberechnungen) mit älteren, bereits überprüften Verfahren zur Erhebung desselben Sachverhalts. Bestenfalls mit dem Goldstandard in diesem Bereich (59).

Zu den Nebengütekriterien lassen sich laut Döring und Bortz (2015) unter anderem Utilität, Akzeptanz, Testfairness und -ökonomie, Transparenz, Unverfälschbarkeit, Zumutbarkeit, Normierung und Vergleichbarkeit zählen (60).

Döring und Bortz schlagen (2015) zusätzlich die Überprüfung der **internen Konsistenz** eines Messinstruments vor (60). Sie soll angeben wie homogen ein Messinstrument ist und die Eindimensionalität belegen. Häufig erfolgt die Angabe über die Berechnung des Cronbachs α -Wertes (61). Der Kennwert Cronbachs α ist jedoch laut Döring und Bortz als internes Konsistenzmaß methodisch kritisch zu betrachten (60). Seine Höhe ist beispielsweise in starkem Maße abhängig von der Anzahl der Testitems (60). Je mehr Fragen oder Items ein Messinstrument besitzt, desto höhere Cronbachs α -Werte werden erreicht.

Neben den Haupt- und Nebengütekriterien ist eine weitere wesentliche Eigenschaft eines Messinstruments die **Sensitivität** gegenüber Veränderungen. Unter dieser Eigenschaft versteht man die Fähigkeit, innerhalb eines Beobachtungszeitraums auftretende Veränderungen zu erheben. Die besondere Bedeutung dieser Eigenschaft wird durch die Tatsache beschrieben, dass positive Veränderungen innerhalb eines

bestimmten Zeitraums das klassische therapeutische Ziel darstellen (62). Die Veränderungssensitivität kann als die Fähigkeit eines Messinstruments definiert werden, auf Änderungen des gemessenen Sachverhalts beziehungsweise Konstrukts zu reagieren, unabhängig davon, ob die gemessene Änderung relevant oder bedeutsam ist (63,64). Um wiederum bestimmen zu können, ob eine gemessene Änderung relevant und klinisch bedeutsam ist, benötigt ein Messinstrument ebenfalls eine hohe **Empfindlichkeit** (engl. „responsiveness“) (63–65). Maße zur Bestimmung der Veränderungssensitivität sind zum einen der Unterschied zwischen Prä- und Postmessung ($T_2 - T_1$) und zum anderen die standardisierte Effektgröße ($(T_2 - T_1) / \text{Standardabweichung}_{T_1}$) sowie der standardisierte Mittelwert der Antworten/Ergebnisse ($(T_2 - T_1) / \text{Standardabweichung}_{T_1 - T_2}$) (66,67). Um beurteilen zu können, ob eine Änderung klinisch bedeutsam ist, muss die „minimal important difference“ berechnet werden. Die Differenz zwischen Prä- und Postmessung muss hierbei größer als 0,3-fache beziehungsweise 0,5-fache die Standardabweichung der Prä-Messung sein (68,69).

5.2 Subjektive Messinstrumente

Unter den subjektiven, patientenbasierten Messinstrumenten (englisch: „patient reported outcome“) ist der „**Disability of the Arm, Shoulder and Hand**“ **Fragebogen (DASH)** der bekannteste und gängigste Fragebogen (58). Der DASH wurde bereits in viele Sprachen übersetzt und validiert. Germann und Kollegen haben im Jahre 2003 eine deutsche Version des DASH validiert und die Ergebnisse veröffentlicht (21). Der DASH erhebt die Funktionsfähigkeit des Arms, der Schulter und der Hand, sowie die Beschwerden des Patienten durch eine Selbstbeurteilung des Patienten. Er ist insgesamt in drei Module unterteilt. Das erste Modul beinhaltet Fragen zur Funktion und Aktivität im Alltag (20 Fragen), zu sozialen Aspekten (2 Fragen), eventuell vorhandenen Symptomen (5 Fragen), Schlafstörungen (1 Frage), zur sexuellen Aktivität (1 Frage) und zum Selbstwertgefühl (1 Frage). Die Schwierigkeiten in der Ausführung der alltäglichen Tätigkeiten, sowie die weiteren Fragen werden anhand einer 5-Punkt-Likertskala von „1 - keine Schwierigkeiten“ bis „5 - nicht möglich“ bewertet. Alle Fragen werden von dem Patienten durch eine Einschätzung, wie er die folgende Tätigkeit in der vergangenen Woche durchgeführt hat bzw. ausüben konnte, beantwortet. Dabei ist es nicht entscheidend, mit welchem Arm oder welcher Hand die

Tätigkeit ausgeübt wurde. Im Falle, dass eine bestimmte Tätigkeit in der vergangenen Woche nicht ausgeführt wurde, sollte die selbst eingeschätzte Fähigkeit zur potenziellen Bewältigung dieser Aufgabe angekreuzt werden. Das zweite und dritte Modul des DASH sind optional und beschäftigen sich mit Kriterien bezüglich Musik und Sport sowie Arbeit und Beruf und umfassen jeweils 4 weitere Fragen.

Für die Bestimmung des DASH-Funktionsfähigkeitswertes der einzelnen Module werden zunächst die Antwortpunkte der Fragen summiert, und somit ein Rohwert gebildet. Der maximal erreichbare Rohwert für das erste Modul beträgt 150 Punkte und das mögliche Minimum liegt bei 30 Punkten. Für Modul 2 und 3 liegen der maximale Rohwert bei jeweils 20 Punkten und der minimale Rohwert bei jeweils 4 Punkten. Im 1. Modul dürfen 3 Fragen unbeantwortet bleiben, bei den Modulen 2 und 3 kann nur ein Ergebnis berechnet werden, wenn alle Fragen beantwortet werden. Der jeweilige Rohwert wird durch die Anzahl der beantworteten Fragen dividiert, davon wird eins subtrahiert und anschließend das Ergebnis mit 25 multipliziert. Durch diese Berechnung werden die Rohwerte der einzelnen Module in einen Funktionsfähigkeitswert von 0 bis 100 Punkten konvertiert, wobei der Wert „0“ keine Einschränkung bedeutete und der Wert „100“ für die höchst mögliche Einschränkung steht (21).

Ein weiterer, vom Patienten durch Selbsteinschätzung ausgefüllter, Fragebogen ist der **„Shoulder Pain and Disability Index“ (SPADI)**. Er erhebt zwei unterschiedliche Dimensionen. Die Schmerzdimension umfasst fünf Fragen zum Schweregrad des Schmerzes des Patienten. Die zweite Dimension umfasst insgesamt acht Fragen zur funktionellen Aktivität der oberen Extremität. Hierbei sollen die Schwierigkeiten des Patienten bei der Durchführung alltäglicher Tätigkeiten mit der oberen Extremität erhoben werden (70). Auch der SPADI-Fragebogen wurde bereits auf seine Reliabilität und Validität hin überprüft und diese wurden in mehreren Studien bestätigt (70). Um die Fragen zu beantworten, sollen die Patienten jeweils auf einer Likert-Skala die Zahl ankreuzen, welche ihren Schmerz bzw. ihre Schwierigkeiten bei der Durchführung einer bestimmten Aufgabe am besten beschreiben. Die Ausprägungen für die Schmerzdimension reichen von „0 - überhaupt keine Schmerzen“ bis hin zu „10 - schlimmste denkbare Schmerzen“, für die funktionellen Aktivitäten von „0 - keine Schwierigkeit“ bis hin zu „10 - so schwierig, dass man Hilfe benötigt“. Die Bewertungen aus beiden Dimensionen werden addiert und durch die Anzahl der maximal möglichen

Punkte (130 Punkte) dividiert. Um die Gesamtpunktzahl des SPADI zu erhalten wird das Ergebnis noch mit 100 multipliziert. Eine Gesamtpunktzahl von „0%“ beschreibt das beste Ergebnis (ein Patient ohne Schmerzen und keinerlei Alltagseinschränkungen), „100%“ beschreiben hingegen das schlechteste Ergebnis (ein Patient mit starken Schmerzen und maximaler Einschränkung und Hilfsbedürftigkeit im Alltag). Die kleinste nachweisbare Veränderung (90%-Konfidenzintervall) beträgt 13 Punkte, eine geringere Änderung kann auf einen Messfehler zurückgeführt werden (70).

5.3 Subjektiv-objektive Messinstrumente

Ein Schulterassessment, welches sowohl subjektive, als auch objektive Elemente enthält, ist der **„Constant-Murley-Score“**. Er setzt sich aus vier Parametern, zwei subjektiven und zwei objektiven, zusammen. Zu den subjektiven Parametern zählen der Schmerz und die Fähigkeit des Patienten, die Alltagsaktivitäten auszuführen, während zu den objektiven Parametern die aktive Beweglichkeit der Schulter und die Schulterkraft zählen. Der Schmerzbeurteilung werden insgesamt 15 Punkte zugeschrieben, wobei die komplette Abwesenheit von Schmerz 15 Punkte ergibt und maximal starke Schmerzen während den Alltagsaktivitäten mit 0 Punkten bewertet wird. Die subjektive Einschätzung, die Alltagsaktivitäten durchführen zu können, wird mit maximal 20 Punkten bewertet. Dieser Parameter umfasst die Beurteilung der Arbeitsfähigkeit, der Ausführung gewohnter Sport- und Freizeitaktivitäten sowie die Bewertung von Schlafstörungen und die Positionierung der Hand zur Durchführung alltäglicher Aktivitäten. Die aktive Beweglichkeit der Schulter wird mit insgesamt 40 Punkten bewertet. Unterteilt wird die Beweglichkeit hinsichtlich der möglichen Bewegungsebenen des Schultergelenks (Flexion/Anteversion 10 Punkte, Abduktion 10 Punkte, Außen- und Innenrotation jeweils 10 Punkte). Die jeweilige Punktezuordnung zu den einzelnen Bewegungsumfängen kann einer entsprechenden Tabelle entnommen werden (22). Die Schulterkraft wird im Sitzen bei einer Abduktion von 90° in der Skapularebene (Abduktion der Schulter mit einem 30-45° Winkel zur Horizontalebene) mit einem Zugmesser gemessen. Als Messpunkt wird das Handgelenk beschrieben (22). Für jedes statisch gehaltene Gewicht von einem englischen Pfund (ca. 0,45 Kilogramm) wird ein Punkt notiert. Maximal sind bei diesem Parameter 25 Punkte zu erreichen. Die Punkte, die den einzelnen Parametern

zugeordnet werden, summieren sich zu einem Gesamtwert von maximal 100 Punkten auf. „0“ Punkte sind mit maximalen Schmerzen und keiner relevanten Schulterfunktion gleichzusetzen, „100“ Punkte beschreiben hingegen einen schmerzfreien Patienten mit voll beweglicher und optimal funktionsfähiger Schulter (22).

5.4 Objektive Messinstrumente

Ein objektives Messverfahren zur Bewertung der Beweglichkeit von Gelenken im Allgemeinen ist die sogenannte **Neutral-Null-Methode**. Im englischsprachigen Raum wäre hier die Rede von einem „clinician reported outcome“, bei welchem die Beurteilung des Zustands des Patienten, hier die Schulterfunktion, durch einen geschulten Rater (medizinische Fachkraft) stattfindet (58). Bei der Neutral-Null-Methode handelt es sich um einen standardisierten orthopädischen Bewertungsindex, welcher das Bewegungsausmaß eines Gelenks, hier beispielsweise des Schultergelenks, in Winkelgraden um eine bestimmte Achse aus der Neutralstellung heraus angibt. Dadurch kann die Beweglichkeit objektiv bestimmt und Einschränkungen anhand von Normwerten erkannt werden (71). Die Untersuchung des Bewegungsausmaßes der Schulter erfolgt sowohl aktiv als auch passiv im Seitenvergleich. Die folgenden Bewegungsrichtungen sind zur Überprüfung der Schulterfunktionalität zu messen: Anteversion und Retroversion, Abduktion, rein glenohumerale Abduktion, Außenrotation des hängenden Arms, sowie Außen- und Innenrotation bei 90° Abduktion in der Scapulaebene (72).

Der **Wolf-Motor-Funktionstest (WMFT)** ist ein weiterer objektiver Test, welcher im Jahre 1989 von Wolf und Kollegen für neurologische Patienten entwickelt wurde (25). In seiner ursprünglichen Version umfasste der WMFT 21 Items und liefert Informationen darüber, bei welchen einfachen und komplexen Bewegungen sowie Alltagsaktivitäten der Patient noch Schwierigkeiten mit der oberen betroffenen Extremität hat. In dieser Version erfolgt die Bewertung rein über die benötigte Zeit zur Durchführung der jeweiligen Aufgaben. 2001 wurde der ursprüngliche WMFT durch Wolf und Kollegen von 21 auf 15 Items gekürzt. Zudem modifizierten Taub und Kollegen (2011) den WMFT. Ihre Version umfasst 17 Items und die Bewertung der einzelnen Aufgaben erfolgt neben der benötigten Zeit, über einen Punktwert, welcher für die funktionelle Fähigkeit vergeben wird (26). Bauder und Kollegen haben 2001 den WMFT dahingehend modifiziert, dass sie noch eine zweite Bewertungsskala

hinzugefügt haben, welche die Bewegungsqualität umfasst. Ihre Version umfasst 19 Items (27). Alle bislang existierenden Versionen des WMFT zielen auf die Beurteilung von neurologischen Patienten mit Hemiparese nach Schlaganfall oder einer traumatisch bedingten Hirnverletzung ab.

6 Fragestellungen

Vor dem Hintergrund der in den vorherigen Kapiteln aufgezeigten Forschungslücke im Bereich der objektiven Messinstrumente für die Schulterfunktion bei unfallchirurgischen und orthopädischen Patienten und der unabdingbaren Notwendigkeit solcher Messinstrumente ist das übergeordnete Ziel der vorliegenden Dissertation die Entwicklung eines Messinstrumentes zur Bestimmung der Schulterfunktion nach orthopädischer Verletzungsdiagnose zur Überprüfung von Therapie-, Interventions- und Rehabilitationserfolgen. Dabei liegt der Fokus auf der Testentwicklung sowie der Überprüfung der Gütekriterien des entwickelten Messinstrumentes anhand einer spezifischen orthopädischen Verletzungsdiagnose - der proximalen Humerusfraktur.

In einem ersten Schritt ging es um die Modifizierung des bislang für neurologische Patienten entwickelten WMFT. Hierbei wurden sowohl die einzelnen Testitems, als auch die Bewertungskriterien auf die speziellen Bedürfnisse und Besonderheiten bei unfallchirurgischen und orthopädischen Patienten angepasst (Manuskript 1). Im Anschluss wurde der modifizierte Wolf-Motor-Funktionstest (WMFT-O) hinsichtlich seiner Objektivität, Reliabilität und internen Konsistenz überprüft. Im Rahmen dieses Vorgehens wurden folgende Fragestellungen bearbeitet:

- ◆ *Ist die Objektivität bei der klinischen Durchführung des angepassten Wolf-Motor-Funktionstests (WMFT-O) für ältere, unfallchirurgische und orthopädische Patienten nach Schulterverletzung gegeben?*
- ◆ *Ist die Reliabilität bei der klinischen Durchführung des angepassten Wolf-Motor-Funktionstests (WMFT-O) für ältere, unfallchirurgische und orthopädische Patienten nach Schulterverletzung gegeben?*
- ◆ *Sind die einzelnen Items des angepassten Wolf-Motor-Funktionstests (WMFT-O) homogen und ist somit die interne Konsistenz des Tests gegeben?*

Vor diesem Hintergrund erfolgte im zweiten Schritt die Anwendung des WMFT-O im Rahmen einer multizentrischen randomisierten, kontrollierten Studie zur Generierung von Datensätzen (Manuskript 2), um anschließend in einem dritten Schritt die Objektivität des videobasierten WMFT-O sowie dessen Veränderungssensitivität zu überprüfen (Manuskript 3). Dabei wurde folgenden Fragestellungen nachgegangen:

- ◆ *Ist die Objektivität der videobasierten Version des WMFT-O für unfallchirurgische Patienten nach proximaler Humerusfraktur vor und nach einer robotergestützten Rehabilitations-maßnahme gegeben?*
- ◆ *Ist der WMFT-O bei unfallchirurgischen Patienten nach proximaler Humerusfraktur vor und nach einer Rehabilitationsmaßnahme gegenüber Veränderungen sensitiv?*
- ◆ *Wenn ja, ist es möglich mit dem WMFT-O bei Patienten nach proximaler Humerusfraktur eine klinisch relevante Veränderung zu messen?*
- ◆ *Sind die klinisch erhobenen Ergebnisse des WMFT-O mit denen der videobasierten Version zu vergleichen?*
- ◆ *Sind die klinisch erhobenen Ergebnisse des WMFT-O mit den Ergebnissen des patientenbasierten Fragebogens (DASH) zu vergleichen?*

Abschließend wurden die Studienergebnisse der in Manuskript 2 vorgestellten multizentrischen randomisierten, kontrollierten Studie analysiert (Manuskript 4). Dabei wurde folgenden Fragestellungen nachgegangen:

- ◆ *Kann eine additive Anwendung des assistiven Robotiktrainings über einen Zeitraum von drei Wochen das funktionelle Outcome bei Patienten mit proximaler Humerusfraktur gegenüber der Kontrollgruppe gemessen am DASH und WMFT-O verbessern?*

7 Testentwicklung und Evaluierung eines Tests zur Bestimmung der Schulterfunktion/ Publikationsübersicht

Im folgenden Kapitel sollen nun die einzelnen Manuskripte zur Testentwicklung und Evaluierung eines objektiven, reliablen und veränderungssensitiven Tests zur Bestimmung der Schulterfunktion bei unfallchirurgischen und orthopädischen Patienten zusammengefasst dargelegt werden. Die Originalmanuskripte befinden sich im Anhang dieser Dissertation.

Die vorliegende Dissertation ist eingebettet in das von der Deutschen Gesetzlichen Unfallversicherung (DGUV) geförderte RASTA-Projekt. Als Reaktion auf den deutlichen Anstieg der Schulter- und Oberarmfrakturen einhergehend mit einem erhöhten Operations- und Rehabilitationsbedarf sowie fehlenden Leitlinien zur optimalen Gestaltung von Trainingskonzepten nach einer solchen Fraktur verfolgt das RASTA-Projekt das übergeordnete Ziel der Untersuchung, ob ein moderner Therapieansatz einem ausschließlich konventionellen physio- und ergotherapeutischen Training überlegen ist. Bei diesem modernen Therapieansatz handelt es sich um ein additives assistives Robotiktraining am Armeo®Spring⁶. Es wird zu einer frühfunktionellen Therapie eingesetzt und soll in Folge dessen zu einer schnelleren Genesung des Patienten sowie zur Vermeidung von anhaltenden Bewegungseinschränkungen des Schultergelenks führen.

Hierzu wurde in nationaler Zusammenarbeit mit den berufsgenossenschaftlichen Unfallkliniken Murnau und Ludwigshafen sowie dem Robert-Bosch-Krankenhaus Stuttgart eine multizentrische, randomisierte, kontrollierte und teil-verblindete Interventionsstudie durchgeführt.

Die vorliegende Dissertation ordnet sich zum einen in die anfängliche Phase des RASTA-Projekts mit der Entwicklung des angepassten Wolf-Motor-Funktionstests (WMFT-O) als sekundärer Outcome-Parameter ein. Zum anderen wurden in der finalen Phase des RASTA-Projekts die durch die randomisierte, kontrollierte Studie gewonnenen Datensätze hinsichtlich der Reliabilität und der Veränderungssensitivität des WMFT-O analysiert. Abschließend wurden die Ergebnisse der randomisierten, kontrollierten Studie analysiert.

⁶ Armeo®Spring: Trainingsroboter der Firma Hocoma zur assistiven Therapie der oberen Extremität zum Wiedererlangen der Arm- und Handfunktion (<https://www.hocoma.com/de/losungen/arm-handtherapie/>)

7.1 Entwicklung des modifizierten Wolf-Motor-Funktionstests (WMFT-O) für ältere Patienten nach Schulterverletzung und klinische Überprüfung der inter- und intra-rater Reliabilität (Manuskript 1)

Oberle C, Becker C, Schölch S, Lenz JU, Studier-Fischer S, Augat P, Schwickert L. Inter-rater and intra-rater reliability of an adapted Wolf motor function test for older patients with shoulder injuries. *Z Gerontol Geriat*, 2018;51(3):293–300.

Hintergrund und Zielsetzung: Bislang fehlen objektive Messinstrumente zur Erhebung der funktionellen Leistungsfähigkeit der Schulter bei unfallchirurgischen Patienten nach Schulterverletzung. Der original WMFT aus dem Jahre 1989 bewertet seither objektiv die Funktion der oberen Extremitäten von Patienten mit neurologischen Störungen und Schädel-Hirn-Verletzungen anhand der benötigten Zeit zur Durchführung 21 bestimmter Aufgaben (25). Nach einer durchgeführten Modifikation von Taub und Kollegen (2011) wird die funktionelle Fähigkeit (26) und nach einer weiteren durchgeführten Modifikation von Bauder und Kollegen (2001) ebenfalls die Bewegungsqualität bewertet (27). Zudem wurde der WMFT auf 17 Items reduziert. Im Rahmen dieser Studie sollen die Bewegungsaufgaben und die Bewertungskriterien des original WMFT auf unfallchirurgische und orthopädische Patienten nach Schulterverletzungen angepasst und die Inter- und Intra-Rater Reliabilität sowie die interne Konsistenz der klinischen Durchführung dieser angepassten Version überprüft werden.

Patienten und Methodik: Im Rahmen der Testentwicklung wurden die Bewertungskriterien, welche bislang auf neurologische Patienten ausgerichtet waren und zum Großteil auf zu beobachtenden Synergismen beruhten, auf das Kollektiv der unfallchirurgischen und orthopädischen Patienten angepasst. Hierzu wurde zunächst der WMFT mit seinen angepassten Bewertungskriterien mit fünf unterschiedlich stark betroffenen Patienten nach Humerusfraktur, welche sich in unterschiedlichen Stadien ihre Genesung befanden, durchgeführt. Die durchgeführten Bewegungsaufgaben wurden jeweils von zwei bis vier unabhängigen Ratern bewertet, um daraus Rückschlüsse auf die Bewertungskriterien ziehen zu können. Neben der Anpassung der Bewertungskriterien wurden zu den bereits existierenden 17 Items drei weitere hinzugefügt. Der hier verwendete ursprüngliche Leitfaden von Bauer und Kollegen

(2001) wurde um ein Item für die Abduktion in der Schulter mit Gewicht erweitert (27). Zusätzlich erhielten die Patienten die Möglichkeit zwei Items mit einer 15cm hohen Kiste durchzuführen, wenn Sie aufgrund ihrer schwachen Abduktionsfähigkeit in der Schulter nicht in der Lage waren, diese mit einer 30cm hohen Kiste durchzuführen. Ergänzend zu den drei hinzugefügten Items wurden drei der ursprünglichen Items modifiziert, um diese noch besser auf das Kollektiv der unfallchirurgischen und orthopädischen Patienten auszurichten. Der so entstandene Wolf-Motor-Funktionstest-Orthopädie (WMFT-O) wurde anschließend bei 20 älteren Patienten (Median 80,5 Jahre alt) mit chirurgisch oder konservativ behandelter Schulterverletzung in einem Test-Re-Test Design (weniger als 2 Tage auseinander) angewendet. Damit zu einem späteren Zeitpunkt der WMFT-O auch verblindet und ortsunabhängig ausgewertet werden kann, wurde dieser auf Video aufgezeichnet. Um die optimale Kameraanzahl und -position herauszufinden, wurden zunächst alle 20 Items des WMFT-O von der Versuchsleiterin ausgeführt und diese aus der frontalen, den beiden seitlichen und vier diagonalen Positionen, sowie von hinten und aus der Vogelperspektive gefilmt. Im Anschluss daran wurden die so entstandenen Videos ausgewertet und zwei optimale Kameraperspektiven identifiziert. Um sowohl die klinischen Rater vor Beginn der Studie, als auch im Verlauf weitere Rater für die videobasierte Auswertung des WMFT-O schulen zu können, wurden zwei Lehrvideos⁷ als Leitfaden erstellt. Zur Bestimmung der Inter-Rater Reliabilität des klinischen WMFT-O bewerteten zwei unabhängige Rater die Leistung der Patienten. Einer der Rater bewertete die Leistung zu einem zweiten Messzeitpunkt erneut um die Intra-Rater Reliabilität zu erheben. Inter- und Intra-Rater Reliabilität wurden durch gewichtete Cohens Kappa-Statistiken mit entsprechenden Konfidenzintervallen überprüft. Zur Bestimmung der internen Konsistenz wurden Cronbachs α -Werte berechnet.

Ergebnisse: Für die Inter-Rater Reliabilität reichten die Cohens Kappa-Werte für die funktionelle Fähigkeit von 0,84-1,00 und für die Bewegungsqualität von 0,79-1,00. Für die Intra-Rater Reliabilität lagen die Werte bei 0,71-1,00. Die Cronbachs α -Werte lagen für die funktionelle Fähigkeit über 0,94 und für die Bewegungsqualität über 0,96.

⁷ Lehrvideo Durchführung: <https://www.youtube.com/watch?v=8yDcWK9Xrtw>
Bewertung: <https://www.youtube.com/watch?v=MsfrGauYHfc>

Diskussion und Schlussfolgerungen: Der WMFT-O zeigt eine gute Intra- und Inter-Rater-Reliabilität und eine hohe interne Konsistenz für unfallchirurgische und orthopädische Patienten nach Schulterverletzungen auf. Er ist somit ein objektiver, praktikabler und zuverlässiger Test zur Ermittlung der Leistungsfähigkeit der oberen Extremität nach unfallchirurgischen und orthopädischen Verletzungen. Um die Generalisierbarkeit zu erhöhen, müssen weitere Untersuchungen mit einem größeren Patientenkollektiv sowie jüngeren Patienten durchgeführt werden. Des Weiteren muss die Veränderungssensitivität des WMFT-O untersucht werden, um diesen in Zukunft zur Verlaufsanalyse in der Rehabilitation und zur Beurteilung von Interventionserfolgen einsetzen zu können.

7.2 Anwendung des modifizierten Wolf-Motor-Funktionstests (WMFT-O) im Rahmen einer multizentrischen randomisierten, kontrollierten und teilverblindeten Studie zur Generierung von Datensätzen zur Überprüfung der Testgütekriterien (Manuskript 2)

Nerz C, Schwickert L, Becker C, Studier-Fischer S, Müßig JA, Augat P. Effectiveness of robot-assisted training added to conventional rehabilitation in patients with humeral fracture early after surgical treatment: protocol of a randomised, controlled, multicentre trial. *Trials*, 2017; 18:589.

Hintergrund und Zielsetzung: Die Inzidenz von Humerusfrakturen nimmt bei Männern ab dem 30. Lebensjahr und bei Frauen ab dem 40. Lebensjahr stark zu (13). Zudem betreffen ab dem 40. Lebensjahr 76% aller Humerusfrakturen den proximalen Anteil des Humerus (42). Die funktionelle Wiederherstellung des Oberarms nach PHF ist langwierig und die Ergebnisse sind oft unbefriedigend. Zudem geht die Rehabilitation mit langen Immobilisationsperioden einher. Evidenzbasierte Leitlinien für die therapeutische Nachbehandlung und funktionelle Wiederherstellung fehlen bislang weitestgehend. Die daraus resultierenden Konsequenzen sind der Verlust der Muskelmasse und der motorischen Leistungsfähigkeit in der betroffenen Extremität sowie eine reduzierte Bewegungsfreiheit. Diesen Verlusten und Einschränkungen soll durch eine gezielte Trainingsintervention mit Hilfe eines additiv eingesetzten, assistiven Trainingsroboters zur Gewichtsentlastung und funktionellen Unterstützung

der betroffenen Extremität, entgegengewirkt werden. Durch diese innovative Intervention soll eine kürzere Immobilisierung erreicht werden. Bislang wurde dieser Ansatz nur in wenigen Studien mit geringen Fallzahlen angewendet (18–20). Ziel ist es daher, zu untersuchen, ob das additive, assistive Robotiktraining am Armeo®Spring begleitend zur konventionellen Ergo- und Physiotherapie dazu beitragen kann, die Leistungsfähigkeit der Schulter im Vergleich zu einer Kontrollgruppe zu verbessern und somit eine schnellere Genesung zu ermöglichen.

Patienten und Methodik: Patienten zwischen 35 und 70 Jahren mit PHF (AO 11) und chirurgischer Nachbehandlung wurden in drei Kliniken in Deutschland (Robert-Bosch-Krankenhaus Stuttgart, Berufsgenossenschaftliche Kliniken Murnau und Ludwigshafen) rekrutiert und in eine Interventions- und eine Kontrollgruppe randomisiert. Ausschlusskriterien waren eine eingeschränkte Kognition (≤ 10 SOMC⁸; (73)), starke Bewegungsschmerzen in der betroffenen Extremität (VAS-Skala⁹ ≤ 5 ; (74)), eine Gehgeschwindigkeit $> 0,8$ m/Sek., chronische Herzerkrankungen (NYHA III-IV¹⁰), chronische Lungenerkrankungen (COPD GOLD III-IV¹¹), und eine unzureichende Sehfähigkeit, sowie Hörfähigkeit. Isolierte Tuberculum majus Frakturen (AO11, A1) sowie Frakturen, welche die Gelenkpfanne, den Plexus- oder Axillarinnerv betreffen und Mehrfachfrakturen wurden ebenfalls ausgeschlossen. Die Teilnehmer wurden zwischen der vierten und siebten postoperativen Woche eingeschlossen und vor und nach der Interventionsphase von drei Wochen, von einem zur Intervention verblindeten Untersucher getestet. Die Hauptendpunkte der Studie umfassten zum einen die patientenorientierte Bewertung der Einschränkung bei alltäglichen Tätigkeiten ausgelöst durch die betroffene Extremität anhand DASH (21), zum anderen die klinischbasierte Messung der motorischen Funktion anhand des WMFT-O (75) sowie die Messung der Bewegungsfreiheit des betroffenen Arms (Goniometer-Messung). Der DASH-Fragebogen wurde zudem prospektiv durch eine postalische Abfrage sechs und zwölf Monate nach Ende der Intervention erhoben. Alle Teilnehmer erhielten, unabhängig vom Studienarm, konventionelle Ergo- und Physiotherapie. Die Interventionsgruppe erhielt zusätzlich über einen Zeitraum von drei Wochen ein assistives Robotiktraining am Armeo®Spring. Das Training wurde an vier Tagen pro

⁸ "Short Orientation-Memory-Concentration Test" (SOMC) (73)

⁹ "visual analogue scale" (VAS) (74)

¹⁰ Klassifikation von chronischen Herzerkrankungen nach der „New York Heart Association“ (NYHA)

¹¹ Chronische obstruktive Lungenerkrankungen (COPD) klassifiziert nach der „Global Initiative for Chronic Obstructive Lung Disease“ (GOLD)

Woche für jeweils 30 Minuten von einer ausgebildeten Ergo- oder Physiotherapeutin durchgeführt. Die Inhalte des Robotiktrainings waren standardisiert und steigerten sich im Laufe der Zeit hinsichtlich des Schwierigkeitsgrads von eindimensionalen bis hin zu dreidimensionalen Bewegungen.

Diskussion und Schlussfolgerungen: Das Studienprotokoll beschreibt eine randomisierte, kontrollierte und teil-verblindete multizentrische Interventionsstudie der Phase-II. Die Ergebnisse dieser Studie sollen der Evaluierung einer innovativen Therapieform zur schnelleren funktionellen Wiederherstellung bei unfallchirurgischen Patienten nach PHF dienen. Die innerhalb dieser Studie gewonnenen Datensätze sollen im Rahmen der Promotion weitergehend analysiert und ausgewertet werden.

7.3 Überprüfung der Veränderungssensitivität und der inter-rater Reliabilität des videobasierten Wolf-Motor-Funktionstests (WMFT-O) sowie Korrelationsanalyse mit dem „Disability of the Arm, Shoulder and Hand Questionnaire“ (DASH) (Manuskript 3)

Nerz C, Schwickert L, Schölch S, Gordt K, Nolte PC, Kröger I, Augat P, Becker C. Inter-rater reliability, sensitivity to change and responsiveness of the orthopaedic Wolf-Motor-Function-Test as functional capacity measure before and after rehabilitation in patients with proximal humeral fractures. *BMC Musculoskeletal Disorders*, 2019, 20:315.

Hintergrund und Zielsetzung: Die Inzidenz von Humerusfrakturen stieg in den letzten Jahren an, so dass pro Jahr mehr als 85.000 Humerusfrakturen registriert werden (1). Dies entspricht 13% der insgesamt registrierten Frakturen in Deutschland (12). Damit einhergehend nahm auch die Inzidenz der PHF im letzten Jahrzehnt um mehr als 30% zu. 45% aller Humerusfrakturen betreffen das proximale Ende, bei Personen über 40 Jahren sind es bereits 76% (42). Dieser Anstieg geht mit einem erhöhten Operations- und Rehabilitationsbedarf einher (8). Allerdings sind die Leitlinien zu operativen vs. nicht-operativen Behandlungen und postoperativen Nachbehandlungen unzureichend und basieren meist auf Expertenmeinungen (11,76). Es ist zwingend notwendig, die Funktionsfähigkeit der oberen Extremität

objektiv zu bewerten, um so den Rehabilitationserfolg und die Leistungsfähigkeit beurteilen und verschiedene Behandlungsmethoden miteinander vergleichen zu können. Dies erfordert neben patientenbasierten Fragebogen, auch klinisch basierte Messinstrumente (23,24), welche valide, reliabel und veränderungssensitiv sein müssen. Diese objektiven Messinstrumente für die obere Extremität fehlen jedoch bislang (11,76). Ein vielversprechendes objektives Messinstrument für neurologische Patienten ist der WMFT (25). Diese Studie soll den auf unfallchirurgische und orthopädische Patienten angepassten WMFT-O hinsichtlich seiner Gütekriterien überprüfen. Es soll die Objektivität des videobasierten WMFT-O anhand seiner Inter-Rater Reliabilität belegt werden. Zusätzlich soll der Zusammenhang zwischen der klinischen und der videobasierten Version des WMFT-O sowie zwischen der videobasierten Version des WMFT-O und dem DASH analysiert werden. Des Weiteren soll die Veränderungssensitivität des WMFT-O und des DASH zur Messung der Funktionsfähigkeit der oberen Extremität vor und nach der Rehabilitation bei Patienten mit PHF analysiert werden. Ergänzend soll untersucht werden, ob der WMFT-O und der DASH in der Lage sind auch klinisch relevante Veränderungen zu messen.

Patienten und Methodik: Um die Inter-Rater Reliabilität des videobasierten WMFT-O zu überprüfen, wurden 56 Patienten (Mittelwert=61,7 ± 14,7 Jahre alt) nach einer operativ behandelten PHF zu zwei unterschiedlichen Zeitpunkten durch drei verblindete Rater bewertet. Die Inter-Rater Reliabilität wurde durch Fleiss Kappa-Statistiken mit dazugehörigen Konfidenzintervallen überprüft. Pearson Korrelationskoeffizienten wurden berechnet, um den Zusammenhang zwischen der klinischen und der videobasierten Version des WMFT-O, sowie zwischen der videobasierten Version des WMFT-O und dem DASH zu analysieren. Zudem wurde von 40 der 56 Patienten die Leistungsfähigkeit der oberen Extremität vor und nach einer dreiwöchigen additiven assistiven robotergestützten Trainingsintervention anhand des DASH und des WMFT-O bewertet. Für diese 40 Patienten (53,8 ± 5 Jahre alt) wurde die Veränderungssensitivität des WMFT-O und des DASH durch Berechnung des „change score“ (CS), der „standardized effect sizes“ (SES) und des „standardized response means“ (SRM) analysiert. Durch die Berechnung der „minimal important difference“ wurde überprüft, ob die Erhebungsinstrumente in der Lage sind, klinisch relevante Veränderungen zu messen.

Ergebnisse: Die Werte der Fleiss Kappa-Statistik für Inter-Rater Reliabilität des videobasierten WMFT-O reichten für die funktionelle Fähigkeit von 0,33-0,66, für die Bewegungsqualität von 0,27-0,54. Die Korrelation zwischen der klinischen und der videobasierten Version des WMFT-O lag für alle drei Rater zwischen $r=0,77$ und $r=0,84$. Die Korrelation zwischen der klinischen Version des WMFT-O und des DASH-Fragebogens war schwach (WMFT-O: funktionelle Fähigkeit $r=-0,27$, Bewegungsqualität $r=-0,32$). Die Veränderungssensitivität war sowohl für den WMFT-O, als auch für den DASH hoch ($CS>14,9$; $SES>0,8$; $SRM>1,2$). Im Vergleich zum DASH war die Veränderungssensitivität des WMFT-O höher. Sowohl der DASH, als auch der WMFT-O sind in der Lage klinisch relevante Veränderungen zu messen.

Diskussion und Schlussfolgerung: Diese Ergebnisse belegen, dass der WMFT-O ein reliables, veränderungssensitives und empfindliches Instrument zur objektiven Messung der Leistungsfähigkeit der Schulter im Verlaufe von Rehabilitationsmaßnahmen nach PHF ist. Darüber hinaus hat sich gezeigt, dass die videobasierte Version des WMFT-O eine Möglichkeit darstellt, die vollständige Verblindung der Rater innerhalb von Studien sicherzustellen.

7.4 Ergebnisse einer multizentrischen randomisierten, kontrollierten und teil-verblindeten Studie zur Überprüfung der Wirksamkeit eines additiven, assistiven Robotiktrainings nach proximaler Humerusfraktur (Manuskript 4)

Kröger I, Nerz C, Schwickert L, Schölch S, Müßig JA, Studier-Fischer S, Nolte PC, Becker C, Augat P. Robot-assisted training after proximal humeral fracture: A randomised controlled multicentre intervention trial. *Clinical Rehabilitation*. 2021;35(2):242-252. (Original submission version)

Hintergrund und Zielsetzung: Die funktionelle Wiederherstellung nach PHF ist oft unbefriedigend (9). Eine verzögerte Genesung und unvollständige Wiederherstellung der Schulterfunktion stellen hierbei nicht nur soziale und wirtschaftliche Belastungen für die Patienten dar, sondern auch wirtschaftliche Belastungen für das Gesundheitssystem und die Gesellschaft. Die Folgen für betroffene Patienten können

von Muskelatrophie bis hin zum Verlust der Arbeitsfähigkeit und damit unter anderem einhergehenden psychosozialen Problemen reichen (9,10). Trotz der Notwendigkeit gibt es bislang keinen evidenzbasierten Konsens für Rehabilitationsansätze nach PHF (18,76). Ein vielversprechender Ansatz aus der Neurologie ist ein assistives, robotergestütztes Training (17). Mögliche Vorteile eines solchen Ansatzes wurden bislang bei unfallchirurgischen Patienten noch nicht untersucht (11,18). Ziel der vorliegenden Studie ist die Präsentation der Ergebnisse der in Manuskript 2 beschriebenen Studie (77). Hierbei soll untersucht werden, ob ein robotergestütztes Training sicher durchzuführen ist, ob es von den Patienten akzeptiert wird und ob durch diesen additiven Therapieansatz die Schulterfunktion im Vergleich zu einer Kontrollgruppe verbessert wird.

Patienten und Methodik: Von Januar 2016 bis Juni 2018 wurden insgesamt 928 Patienten untersucht. Davon konnten 48 Patienten mit PHF (AO11) (Mittelwert=55,1 Jahre alt) einer von zwei Gruppen zufällig zugeordnet werden: der Kontrollgruppe, welche über einen Zeitraum von drei Wochen konventionelle Ergo- und Physiotherapie erhielt oder der Interventionsgruppe, welche zusätzlich 12 robotergestützte Therapieeinheiten à 30 Minuten erhielt. Die Schulterfunktion wurde vor und nach der Intervention anhand des DASH, dem WMFT-O, der Messung der Bewegungsfreiheit des betroffenen Arms und der Handkraft ermittelt. Der DASH wurde zudem prospektiv durch eine postalische Abfrage 6 und 12 Monate nach Ende der Intervention erhoben.

Ergebnisse: Beide Gruppen verbesserten sich stetig über die drei Messzeitpunkte hinweg. Jedoch zeigte keiner der gemessenen Parameter einen zusätzlichen Nutzen der Robotertherapie auf.

Diskussion und Schlussfolgerung: Die Studie weist einige Limitationen auf. Es handelt sich um eine sehr kleine und zudem, hinsichtlich der operativen Versorgungsart, eine sehr heterogene Studienpopulation. Die Altersobergrenze von 70 Jahren wurde auf der Grundlage des Fördersponsors festgelegt. Die Ergebnisse können daher für eine ältere Population, welche häufig von PHF betroffen ist, unterschiedlich sein. Basierend auf der sehr guten Akzeptanz der Robotertherapie und der sicheren Durchführung dieser, sollten zukünftige Studien untersuchen, ob die konventionelle Therapie durch eine robotergestützte Therapie ersetzt werden kann. Des Weiteren sollten zukünftige Studien den Vorteil der assistiven Robotertherapie nutzen und die Auswirkungen eines früheren Therapiebeginns sowie einer Therapie

mit höheren Intensität, welche über einen längeren Zeitraum durchgeführt wird, untersuchen.

8 Einordnung der Studienergebnisse in den Forschungszusammenhang

Die vorliegenden Studienergebnisse belegen die Eignung des modifizierten Wolf-Motor-Funktionstests-Orthopädie (WMFT-O) als objektives, reliables und gegenüber Veränderungen sensitives Messinstrument zur Erhebung der Schulterfunktion bei unfallchirurgischen und orthopädischen Patienten aller Altersklassen (35-94 Jahre) nach proximaler Humerusfraktur (PHF). Zudem ist der WMFT-O in der Lage, klinisch relevante Veränderungen zu messen.

Durch die in Manuskript 1 beschriebene Modifikation der Testitems des WMFT-O und die Anpassung der Bewertungskriterien der funktionellen Fähigkeit und der Bewegungsqualität wurde eine Möglichkeit für unfallchirurgische und orthopädische Patienten geschaffen, die motorische Leistungsfähigkeit der Schulter auf objektive Weise zu messen. Dies war, laut aktueller Studienlage, bislang nur in einem sehr begrenzten Rahmen möglich (vgl. Kapitel 5.1). Neben dem bislang als Goldstandard geltenden und häufig angewendeten subjektiven und patientenorientierten DASH Fragebogen („patient reported outcome“), gibt es in diesem Bereich weitestgehend nur Messinstrumente, welche sowohl subjektive als auch objektive Elemente enthalten (vgl. Constant-Murley-Score; Kapitel 3.3). Als Ergebnis einer durchgeführten Literaturrecherche im Bereich der objektiven Messinstrumente („clinician reported outcome“) für die funktionelle Leistungsfähigkeit der oberen Extremität konnten lediglich die Neutral-Null-Methode und der WMFT gefunden werden (vgl. Kapitel 3.4). Die Neutral-Null-Methode wird bereits bei unfallchirurgischen und orthopädischen Patienten eingesetzt (18). Ermittelt wird jedoch lediglich die maximal mögliche Bewegungsweite der betroffenen Extremität und nicht die funktionelle Fähigkeit die betroffene Extremität im Alltag einzusetzen. Der vielversprechende Ansatz des WMFT für neurologische Patienten wurde in der vorliegenden Dissertation als Ausgangspunkt zur weiteren Testentwicklung gewählt, um dadurch die aufgezeigte Forschungslücke im Bereich der objektiven Messinstrumente bei orthopädischen Verletzungsdiagnosen zu schließen (vgl. Kapitel 3).

Die Überprüfung der Gütekriterien des modifizierten WMFT-O, welche in Manuskript 1 dargelegt wird, belegt eine hohe Objektivität des klinisch durchgeführten WMFT-O (vgl. Kapitel 5.1). Die sehr guten Ergebnisse der Inter-rater Reliabilität des WMFT-O stehen den bisherigen Ergebnissen des WMFT, welche in der Literatur zu finden sind, nicht

nach. So konnten Pereira und Kollegen (2011) für die brasilianische Version des WMFT (78) und Nijland und Kollegen (2010) für den WMFT innerhalb einer durchgeführten Studie bei Schlaganfallpatienten vergleichbare Werte der Inter-rater Reliabilität aufzeigen (79). Vergleicht man die Ergebnisse hinsichtlich der Objektivität des klinischen WMFT-O mit denen der Originalversion, so kann festgestellt werden, dass Morris und Kollegen (2001) neun Aufgaben identifiziert haben, welche eine niedrige Inter-rater Reliabilität aufweisen (80). Dies ist in der in Manuskript 1 vorgestellten Studie nicht der Fall gewesen (vgl. Kapitel 5.1). Die verbesserte Inter-rater Reliabilität des WMFT-O ist vermutlich auf die durchgeführten Modifikationen zurückzuführen und lässt die Schlussfolgerung zu, dass die auf unfallchirurgische und orthopädische Patienten angepasste Version des WMFT-O zur objektiven Messung der Schulterfunktion geeignet ist.

Betrachtet man die Reliabilität des WMFT-O so lässt sich auch hier feststellen, dass die im Manuskript 1 beschriebenen Ergebnisse denen der Literatur entsprechen. Pereira und Kollegen (2011) analysierten den WMFT-O bei Patienten mit Hemiparese in Brasilien und erhielten weitgehend übereinstimmende Ergebnisse (vgl. Kapitel 5.1) (78).

In der in Manuskript 1 vorgestellten Studie wurde des Weiteren, wie von Wolf und Kollegen (1989) und Taub und Kollegen (2011) empfohlen, die benötigte Zeit zur Ausführung einer bestimmten Aufgabe gemessen (25,26). Konnte eine bestimmte Bewegungsaufgabe nicht durchgeführt werden, so wurde eine maximale Zeit von 120 Sekunden für die Bewegungsaufgabe angenommen. Dieses Vorgehen ist bei Schlaganfallpatienten oder Patienten mit Hemiparese durchaus sinnvoll, bei unfallchirurgischen und orthopädischen Patienten hingegen nicht zwangsläufig. Aufgrund der deutlich verkürzten Initialisierungsphase zur Einleitung einer Bewegung und einer raschen Durchführung der Bewegungsaufgabe bei unfallchirurgischen und orthopädischen Patienten, kommt es zu erheblichen Messfehlern bei der Zeitmessung mit einer Stoppuhr. Daher wird vorgeschlagen für zukünftige Untersuchungen des WMFT-O die benötigte Zeit zur Durchführung der Aufgaben nicht weiter in die Bewertung miteinfließen zu lassen (vgl. Kapitel 5.1).

Auch auf die hier beschriebenen Sensitivitätsprobleme bei der Zeitmessung der durchgeführten Bewegungsaufgaben wird in der Literatur bereits hingewiesen. Als ein Beispiel wird die Aufstehbewegung von einem Stuhl bei dem „Chair raising Test“

beschrieben. Schwickert und Kollegen (2011) empfehlen zur Lösung des Sensitivitätsproblems daher eine Erhöhung der Übungswiederholungsanzahl (18). Dies kann jedoch zu Lern- und Muskelermüdungseffekten führen, welche wiederum die Testergebnisse verzerren. Daher sollte diese Möglichkeit beim WMFT-O nicht als Lösungsansatz in Betracht gezogen werden (vgl. Kapitel 5.1).

Um auf das Thema der Muskelermüdung und die dadurch eventuell resultierenden Verzerrungen der Testergebnisse noch einmal einzugehen, müssen die gesamte Länge des WMFT-O (Gesamtdauer circa 40 Minuten) sowie die einzelnen Bewegungsaufgaben betrachtet werden. Aufgrund von insgesamt 18 Bewegungsaufgaben, davon zwei mit zusätzlichem Gewicht in Form einer Gewichtsmanschette, einer Aufgabe mit einem Zusatzgewicht in einem Eimer und einer mit zusätzlichem Gewicht in Form einer Getränkedose (333g) ermüden die Muskeln bei der Durchführung des WMFT-O. Erhöhtes kompensatorisches Bewegungsverhalten der Patienten im Schulterbereich könnte das Ergebnis sein, welches folglich zu einer niedrigeren Bewertung der Schulterfunktion führt. Eine Lösung hierfür könnte die Entwicklung einer Kurzversion des WMFT-O sein, welche weniger redundant Bewegungsaufgaben enthält und dadurch die Gesamtbelastung für den Patienten reduziert. Ein Versuch hierzu wurde bereits in der Literatur vorgestellt. Bogard und Kollegen (2009) zeigten, dass der original WMFT bei neurologischen Patienten von 17 auf sechs Bewegungsaufgaben gekürzt werden kann (81). Wu und Kollegen (2011) zeigten weiterhin, dass die Ergebnisse der Kurzversion des WMFT mit der der Langversion vergleichbar sind (82). Diese Ergebnisse können zum Anlass genommen werden, um über den berechtigten Nutzen einer verkürzten WMFT-O Version für unfallchirurgische und orthopädische Patienten nachzudenken, insbesondere, da diese Zielgruppe tendenziell stärker unter Ermüdungserscheinungen während der Durchführung des WMFT-O leidet (vgl. Kapitel 5.1).

Versucht man die in Manuskript 3 dargelegten Ergebnisse in den Forschungszusammenhang einzuordnen, so muss man feststellen, dass die Ergebnisse des videobasierten WMFT-O hinsichtlich seiner Objektivität etwas niedriger sind, als die des klinisch durchgeführten WMFT-O (vgl. Kapitel und Kapitel 5.3). Die geringere Übereinstimmung der auf Video aufgenommenen Messung des WMFT-O kann mehrere Gründe haben. Möglicherweise reichte das Training der Video-Rater nicht gänzlich aus. Eine weitere Erklärung für die geringere

Übereinstimmung könnte auch eine unzureichende Positionierung der ausgewählten Videokameras sein und eine oder mehrere andere Kamerapositionen könnte zu besseren Ergebnissen führen. Diese Erkenntnisse sind in weiterführenden Studien zu beachten. Nichtsdestotrotz belegen die in Manuskript 3 dargelegten Ergebnisse nach den Standards von Evans (1996) eine sehr starke Korrelation zwischen der klinisch durchgeführten Messung des WMFT-O und der videobasierten Bewertung (vgl. Kapitel 5.3) (83).

Zur Beurteilung der Behandlungseffekte von Rehabilitationsmaßnahmen und Interventionen bei Erkrankungen des Bewegungsapparates müssen sowohl die funktionelle Fähigkeit sowie die Aktivität bewertet werden (23). Daher sollte der WMFT-O durch weitere Messmethoden ergänzt werden, um ergänzend sowohl Einschränkungen im Alltag, als auch die Aktivität und Partizipation von Patienten bewerten zu können. Eine hierzu weit verbreitete und bereits etablierte Methode ist der DASH Fragebogen. Es wird erwartet, dass die Korrelation zwischen Aktivitätsniveau und funktioneller Fähigkeit häufig niedrig ausfällt (84,85). Daher verwundert es nicht, dass die in Manuskript 3 berechnete Korrelation zwischen dem klinisch basierten WMFT-O (Bewertung der funktionellen Fähigkeit und der Bewegungsqualität) und dem patientenbasierten DASH Fragebogen (Aktivitätsniveau) schwach ausfällt (vgl. Kapitel 5.3). Begründungen hierfür könnten unter anderem sein, dass Patienten aufgrund von psychologischen Problemen, wie beispielsweise ungenügender Selbstwirksamkeit, nicht zu ihrem Aktivitätsniveau zurückkehren. Ein weiterer Erklärungsansatz sind methodische Probleme bei der Durchführung des DASH Fragebogens. Der DASH Fragebogen ist, konträr zum WMFT-O, nicht explizit für die betroffene Extremität konzipiert. Dies bedeutet, dass in Fällen, in denen Aufgaben mit der dominanten Hand ausgeführt werden (z. B. Umdrehen eines Schlüssels in einem Schloss), die Einschränkung der nicht dominanten Extremität nicht durch den DASH Fragebogen erfasst wird. Dies ist nur dann der Fall, wenn die betroffene Extremität auch die dominante ist. Diese Fehleinschätzung könnte vermieden werden, wenn darauf geachtet wird, dass bei der Beantwortung der Fragen des DASH Fragebogens die Beurteilung der Einschränkung immer auf die Durchführung der Tätigkeit mit der betroffenen Extremität bezogen wird. Wenn es nicht möglich ist, die Aktivität des täglichen Lebens mit der betroffenen Extremität auszuführen, muss sich der Patient die Durchführung der Aktivitäten und die dabei möglicherweise auftretenden Einschränkungen so gut wie möglich vorstellen. Die hier

beschriebene Berechnung der Korrelation zwischen einem neuen Testverfahren (WMFT-O) und einem etablierten Messinstrument (DASH) zielt zudem auf die Frage der Validität des WMFT-O ab (59). Misst der WMFT-O auch den Sachverhalt, den er vorgibt zu messen? Aufgrund der schwachen Korrelation zwischen dem WMFT-O und dem DASH empfiehlt sich eine weitere Überprüfung der Validität des WMFT-O. Ein Vergleich des WMFT-O und einer Bewegungsanalyse anhand des Nexus VICON-Systems¹² wäre eine Möglichkeit in einer zukünftigen Studie, die Validität des WMFT-O zu überprüfen (86).

Neben dem Vergleich des WMFT-O und des DASH Fragebogens wurde in der in Manuskript 3 vorgestellten Studie (vgl. Kapitel 5.3) des Weiteren die Veränderungssensitivität des WMFT-O und des DASH Fragebogens untersucht. Der WMFT-O weist sowohl für die funktionelle Fähigkeit als auch für die Bewegungsqualität eine hohe Effektgröße auf und kann daher, gemäß den veröffentlichten Standards Husted und Sawilowsky, als sensitiv gegenüber Veränderungen angesehen werden (87,88). Dieses Ergebnis wurde durch die absoluten Werte der „Standardized effect size“ bestätigt. Dies ist ein relevantes Ergebnis im Hinblick auf die klinische Anwendung, um eine objektive und gegenüber Veränderungen empfindliche Beurteilung der motorischen Funktionsfähigkeit der oberen Extremitäten für unfallchirurgische und orthopädische Patienten in der Rehabilitation bereitzustellen und kann die derzeit überwiegend subjektiven Beurteilungen bereichernd ergänzen. Im Vergleich zum WMFT-O war die Veränderungssensitivität des DASH Fragebogens etwas geringer. Schaut man diesbezüglich in die Literatur, so stößt man auf Ergebnisse von Mac Dermid und Kollegen (2001) und Westphal und Kollegen (2007) zur Überprüfung der Veränderungssensitivität des DASH Fragebogens bei Patienten nach Handgelenksfrakturen (24,89). Sie fanden höhere Werte für die Effektgröße des DASH Fragebogen nach 0-3 Monaten. Nach 3-12 Monaten waren sie etwas niedriger. Dies deutet darauf hin, dass die „Standardized effect size“ und der „Standardized response mean“ vom Zeitpunkt der Prä- und Postmessung abhängen. In den ersten 12 Wochen nach Studienbeginn ist mit einem großen Behandlungs- und Genesungseffekt zu rechnen. Innerhalb der nächsten 40 Wochen kann es zu weiteren Verbesserungen kommen, die Effektgrößen werden jedoch geringer sein (vgl. Kapitel 5.3). In der in

¹² Das Nexus VICON-System ist ein markerbasiertes, dreidimensionales Messsystem zur Bewegungsanalyse

Manuskript 2 beschriebenen Studie wurde die Prä-Messung ungefähr einen Monat nach der Operation durchgeführt. Um herauszufinden, ob der WMFT-O auch zur Beurteilung langfristiger Therapieergebnisse verwendet werden kann, sollten zukünftige Studien durchgeführt werden, in denen die Funktionsfähigkeit der oberen Extremität anhand des WMFT-O nach drei Monaten bewertet wird.

Nachdem in Manuskript 3 die Veränderungssensitivität des WMFT-O und des DASH bei Patienten mit PHF belegt werden konnte, wurde zudem untersucht, ob die gemessenen Veränderungen auch eine klinische Relevanz aufweisen (vgl. Kapitel 5.3). Hierbei wird untersucht, ob die gemessene Veränderung auf einem Zufall oder einen Messfehler zurückzuführen ist oder ob es sich um eine klinisch relevante Verbesserung handelt (64,65). Nur letzteres beschreibt das klassische therapeutische Ziel. In der in Manuskript 3 vorgestellten Studie, konnten sowohl anhand des WMFT-O als auch anhand des DASH klinisch relevante Ergebnisse gemessen werden (vgl. Kapitel 5.3).

Zusammenfassend lässt sich sagen, dass sowohl der WMFT-O, als auch der DASH-Fragebogen veränderungssensitive und komplementäre Messinstrumente sind, anhand welcher die funktionelle Veränderung bei unfallchirurgischen und orthopädischen Patienten nach PHF gemessen werden kann (vgl. Kapitel 5.3).

Aufgrund der oben bereits erwähnten Tatsache, dass zur Beurteilung der Behandlungseffekte von Rehabilitationsmaßnahmen sowohl die funktionelle Fähigkeit, als auch die Aktivität bewertet werden müssen (23), und der bislang mangelnden Existenz klinischbasierter Messinstrumente im Bereich der unfallchirurgischen Schulterverletzungen (vgl. Kapitel 3), wurde der modifizierte WMFT-O bereits in einer multizentrischen, randomisierten, kontrollierten und teil-verblindeten Studie eingesetzt. Er wurde neben dem primären Endpunkt (DASH) zur Überprüfung der Wirksamkeit eines additive, assistiven Robotiktrainings nach PHF als sekundärer Endpunkt eingesetzt (vgl. Kapitel 5.4). Es hat sich gezeigt, dass der WMFT-O auch in einer jüngeren Zielgruppe anwendbar und sensitiv genug ist, um den Rehabilitationserfolg über mehrere Wochen hinweg zu messen. Auch wenn in der in Manuskript 4 ausgewerteten Studie kein zusätzlicher Nutzen der additiven Robotertherapie in der Interventionsgruppe nachgewiesen werden konnte, wurde dennoch eine stetige Verbesserung über alle drei Messzeitpunkte hinweg erfasst (vgl. Kapitel 5.4). Auch hier konnte nur eine schwache bis moderate Korrelation zwischen

den Ergebnissen des DASH und des WMFT-O nachgewiesen werden (vgl. Kapitel 5.4). Wie bereits beschrieben ist eine niedrige Korrelation zwischen dem Aktivitätsniveau der Patienten (gemessen durch den DASH) und der funktionellen Fähigkeit (erhoben anhand des WMFT-O) zu erwarten (84,85). Die Ergebnisse dieser Studie untermauern die Tatsache, dass zur Erhebung der Kapazität und Aktivität sowohl patientenbasierte, als auch klinischbasierte Messinstrumente benötigt werden (23). Des Weiteren hat sich in der in Manuskript 4 vorgestellten Studie gezeigt, dass es sinnvoll ist in zukünftigen multizentrischen, randomisierten, kontrollierten Studien auf den speziell für solche Studien entwickelten videobasierten WMFT-O zurückzugreifen, um zum einen die Verblindung der Rater aufrecht zu erhalten und zum anderen eine zentralisierte Auswertung in einem der Zentren durchzuführen (vgl. Kapitel 5.3). Alles in allem wurde in der in Manuskript 4 vorgestellten Studie der WMFT-O zum ersten Mal zu Zweck eingesetzt, zu welchem er unter anderem entwickelt wurde: Zur Evaluation und anschließenden Verbesserung von Rehabilitationsmaßnahmen nach PHF.

9 Schlussfolgerung und Ausblick

Die in dieser Dissertation festgehaltenen Ergebnisse belegen, dass der hier entwickelte Wolf-Motor-Funktionstest-Orthopädie (WMFT-O) ein wertvolles, bislang noch nicht dagewesenes objektives, reliables und veränderungssensitives Messinstrument für die Erfassung der Leistungsfähigkeit von unfallchirurgischen und orthopädischen Patienten nach proximaler Humerusfraktur (PHF) ist.

Die Objektivität der klinisch durchgeführten Version des WMFT-O wurde sowohl für die Bewertung der funktionellen Fähigkeit, als auch für die Bewertung der Bewegungsqualität mittels der Überprüfung der Inter-Rater-Reliabilität belegt. Die Objektivität der klinischen Version des WMFT-O ist über alle Testitems hinweg etwas höher als die der videobasierten Version des WMFT-O. Daher werden, vor der Anwendung der videobasierten Version des WMFT-O in Folgestudien, kleine Anpassungen der Kamerapositionen zur Verbesserung der Objektivität empfohlen. Nichtsdestotrotz ist die videobasierte Version des WMFT-O eine gute Möglichkeit eine Vollverblindung innerhalb von Interventionsstudien zu ermöglichen, welche sonst aufgrund der Kontamination von Therapeuten und Beobachtern in den meisten Fällen nicht vollständig gegeben ist. Die Reliabilität der klinischen Version des WMFT-O wurde anhand der Intra-Rater-Reliabilität überprüft und kann als gegeben angesehen werden. Angesichts der gegebenen Objektivität und der hohen Reliabilität wurde der WMFT-O bereits als zweiter Outcome-Parameter in einer multizentrischen, randomisierten, kontrollierten und teil-verblindeten Studie zur Erhebung der Schulterfunktion eingesetzt (90). Um mögliche Verbesserungen in Folge einer rehabilitativen Anwendung messen zu können, wurde der modifizierte WMFT-O auch auf seine Veränderungssensitivität hin überprüft. Die Ergebnisse zeugen von einer hohen Sensitivität des WMFT-O gegenüber Veränderungen im zeitlichen Verlauf einer Therapie. Ergänzend ist der WMFT-O in der Lage, klinisch relevante Verbesserungen zu messen, welche in der therapeutischen Nachbehandlung von Verletzungen das klassische therapeutische Ziel beschreiben.

Betrachtet man den im Rahmen dieser Dissertation modifizierten und auf unfallchirurgische und orthopädische Patienten angepassten WMFT-O im Vergleich zum bislang gängig eingesetzten und subjektiven DASH Fragebogen, so fällt auf,

dass die Ergebnisse dieser beiden Messinstrumente nur schwach miteinander korrelieren. Diese Tatsache unterstützt die Aussage von Vincent und Kollegen (2013), dass man zur Messung der Kapazität und Aktivität sowohl subjektive, patientenbasierende Fragebögen, als auch objektive, klinischbasierte Messinstrumente benötigt (23). Der WMFT-O schließt die hier dargestellte Forschungslücke im Bereich der objektiven Messinstrumente für die Schulterfunktion bei unfallchirurgischen und orthopädischen Patienten aller Altersklassen (35-94 Jahre alt) mit PHF. Um die Generalisierbarkeit des WMFT-O weiter zu erhöhen, sollten Studien an Patienten mit weiteren orthopädischen Diagnosen der oberen Extremität durchgeführt werden.

Neben dem Einsatz des WMFT-O als Outcome-Parameter in klinischen Studien, ist es durch seine Anwendung möglich therapeutischen Fortschritten in der Nachbehandlung von PHF zu bewerten und daraus resultierend Leitlinien für die therapeutische Nachbehandlung anzupassen. Unter Zuhilfenahme des WMFT-O können so neue, innovative Therapieansätze leichter beurteilt und bereits existierende Ansätze optimiert werden. Dies stellt, nicht nur vor dem Hintergrund des demografischen Wandels und der damit immer häufiger auftretenden Diagnose der PHF, eine wichtige Daseinsberechtigung für den WMFT-O dar. Einheitliche und insbesondere optimierte Leitlinien zum Nachbehandlungsschema von PHF fördern nicht nur die schnelle Genesung, sondern sie verbessern auch das allgemeine physische und das damit einhergehende psychische Wohlbefinden von Patienten nach einer solchen Fraktur. Zudem können diese somit zur Reduktion der negativen Konsequenzen, wie beispielsweise der andauernden Arbeitsunfähigkeit, der Pflegebedürftigkeit und früheren Sterblichkeit und den daraus resultierenden Kosten für das Gesundheitssystem, beitragen.

In einem nächsten Schritt könnte nun die Entwicklung einer anwenderfreundlichen Kurzversion des WMFT-O in den Mittelpunkt rücken. Diese würde aufgrund von Zeitersparnis bei der Durchführung und Bewertung sowie geringeren Konzentrations- und Ermüdungseffekten die Einsatzmöglichkeiten des WMFT-O im klinischen Alltag weiter erhöhen. Zudem wäre es interessant, Studien mit

Inertialsensoren¹³ durchzuführen, umso weitere Erkenntnisse über den funktionellen Einsatz der betroffenen Extremität im Alltag der Patienten zu erhalten und die Ergebnisse des DASH zur Aktivität des Patienten um eine objektive Messfunktion zu ergänzen.

¹³ Inertialsensoren dienen der Messung translatorischer und rotatorischer Beschleunigungen. Kombiniert man mehrere Inertialsensoren miteinander erhält man eine inertielle Messeinheit, mit welcher Beschleunigungen der sechs Freiheitsgraden gemessen werden können.

10 Literaturverzeichnis

1. © Statistisches Bundesamt. Krankenhausstatistik - Diagnosedaten der Patienten und Patientinnen in Krankenhäusern [Internet]. 2017 [zitiert 25. Februar 2019]. Verfügbar unter:
<https://www-genesis.destatis.de/genesis/online?sequenz=statistikTabellen&selectionname=23131>
2. Statistisches Bundesamt [Destatis] - Zentraler Auskunftsdienst. Krankenhausstatistik - Diagnosedaten der Patienten und Patientinnen in Krankenhäusern [Internet]. Destatis. 2020 [zitiert 7. Juli 2020]. Verfügbar unter: http://www.gbe-bund.de/oowa921-install/servlet/oowa/aw92/WS0100/_XWD_FORMPROC?TARGET=&PAGE=_XWD_D_112&OPINDEX=2&HANDLER=_XWD_CUBE.SETPGS&DATACUBE=_XWD_140&D.001=1000001&D.946=17898&D.011=44302#SOURCES
3. Kim SH, Szabo RM, Marder RA. Epidemiology of humerus fractures in the United States: nationwide emergency department sample, 2008. *Arthritis Care & Research*. 2012;64(3):407–14.
4. Chu SP, Kelsey JL, Keegan THM, Sternfeld B, Prill M, Quesenberry CP, u. a. Risk factors for proximal humerus fracture. *Am J Epidemiol*. 15. August 2004;160(4):360–7.
5. Nelißen W. Funktionsdiagnostik nach verschieden operativ versorgten Subkapitalen Humerus- und Humerusschaftfrakturen. [Bochum]: Ruhr-Universität Bochum; 2006.
6. Frey S. Rationelle Therapie von subkapitalen Humerusfrakturen und Humeruskopffrakturen [doctor medicinae dentium]. [Münster]: Westfälischen Wilhelms-Universität Münster; 2003.
7. Lill H, Blum J, Böhm D, Boilaueu P. Die proximale Humerusfraktur: Neue Techniken, neue Implantate - Tipps und Tricks. 1., Aufl. Stuttgart: Thieme; 2006. 212 S.
8. Häussler B, Gothe H, Göl D, Glaeske G, Pientka L, Felsenberg D. Epidemiology, treatment and costs of osteoporosis in Germany—the BoneEVA Study. *Osteoporos Int*. 1. Januar 2007;18(1):77–84.
9. Einsiedel T, Becker C, Stengel D, Schmelz A, Kramer M, Däxle M, u. a. [Do injuries of the upper extremity in geriatric patients end up in helplessness? A prospective study for the outcome of distal radius and proximal humerus fractures in individuals over 65]. *Z Gerontol Geriatr*. Dezember 2006;39(6):451–61.
10. Hodgson S. Proximal Humerus Fracture Rehabilitation. *Clinical Orthopaedics and Related Research (1976-2007)*. Januar 2006;442:131–138.
11. Handoll HHG, Brorson S. Interventions for treating proximal humeral fractures in adults. *Cochrane Database Syst Rev*. 11. November 2015;(11):CD000434.

12. Benzinger P, Riem S, Bauer J, Jaensch A, Becker C, Büchele G, u. a. Risk of institutionalization following fragility fractures in older people. *Osteoporos Int.* 2019;in press.
13. Court-Brown CM, Garg A, McQueen MM. The epidemiology of proximal humeral fractures. *Acta Orthop Scand.* August 2001;72(4):365–71.
14. Passaretti D, Candela V, Sessa P, Gumina S. Epidemiology of proximal humeral fractures: a detailed survey of 711 patients in a metropolitan area. *Journal of Shoulder and Elbow Surgery.* 1. Dezember 2017;26(12):2117–24.
15. Sudhues H. Outcome-Analyse proximaler Oberarmfrakturen und distaler Radiusfrakturen bei geriatrischen Patienten. 2009;79.
16. Kutscha-Lissberg F, Kollig E, Keller M, Muhr G. [Fracture management in the aged]. *Orthopade.* April 2000;29(4):274–80.
17. Veerbeek JM, Langbroek-Amersfoort AC, van Wegen EEH, Meskers CGM, Kwakkel G. Effects of Robot-Assisted Therapy for the Upper Limb After Stroke. *Neurorehabil Neural Repair.* 2017;31(2):107–21.
18. Schwickert L, Klenk J, Stähler A, Becker C, Lindemann U. Robotic-assisted rehabilitation of proximal humerus fractures in virtual environments: a pilot study. *Z Gerontol Geriatr.* Dezember 2011;44(6):387–92.
19. Liao W-W, Wu C-Y, Hsieh Y-W, Lin K-C, Chang W-Y. Effects of robot-assisted upper limb rehabilitation on daily function and real-world arm activity in patients with chronic stroke: a randomized controlled trial. *Clin Rehabil.* Februar 2012;26(2):111–20.
20. Padilla-Castaneda MA, Sotgiu E, Frisoli A, Bergamasco M, Orsini P, Martiradonna A, u. a. A virtual reality system for robotic-assisted orthopedic rehabilitation of forearm and elbow fractures. In: 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems [Internet]. Tokyo: IEEE; 2013 [zitiert 7. April 2020]. S. 1506–11. Verfügbar unter: <http://ieeexplore.ieee.org/document/6696548/>
21. Germann G, Harth A, Wind G, Demir E. Standardisierung und Validierung der deutschen Version 2.0 des “Disability of Arm, Shoulder, Hand” (DASH)-Fragebogens zur Outcome-Messung an der oberen Extremität. *Unfallchirurg.* 1. Januar 2003;106(1):13–9.
22. Constant CR, Murley AH. A clinical method of functional assessment of the shoulder. *Clin Orthop Relat Res.* Januar 1987;(214):160–4.
23. Vincent JI, MacDermid JC, King GJ, Grewal R. Validity and sensitivity to change of patient-reported pain and disability measures for elbow pathologies. *J Orthop Sports Phys Ther.* April 2013;43(4):263–74.
24. Westphal T. [Reliability and responsiveness of the German version of the Disabilities of the Arm, Shoulder and Hand questionnaire (DASH)]. *Unfallchirurg.* Juni 2007;110(6):548–52.

25. Wolf SL, Lecraw DE, Barton LA, Jann BB. Forced use of hemiplegic upper extremities to reverse the effect of learned nonuse among chronic stroke and head-injured patients. *Exp Neurol.* Mai 1989;104(2):125–32.
26. Taub E, Morris DM, Cargo J. Wolf Motor Function Test (WMFT) Manual [Internet]. Birmingham: University of Alabama: CI Therapy Research Group; 2011 [zitiert 7. April 2020]. Verfügbar unter:
https://www.uab.edu/citherapy/images/pdf_files/CIT_Training_WMFT_Manual.pdf
27. Bauder H, Taub E, Miltner W. Behandlung motorischer Störungen nach Schlaganfall: Die Taubsche Bewegungsinduktionstherapie. Göttingen: Hogrefe; 2001. 135 S.
28. Raschke MJ, Stange R. Alterstraumatologie: Prophylaxe, Therapie und Rehabilitation - mit Zugang zum Elsevier-Portal. München Jena: Urban & Fischer Verlag/Elsevier GmbH; 2008. 576 S.
29. Burkhart KJ, Dietz SO, Bastian L, Thelen U, Hoffmann R, Müller LP. The treatment of proximal humeral fracture in adults. *Dtsch Arztebl Int.* September 2013;110(35–36):591–7.
30. Szyszkowitz R, Schippinger G. [Fractures of the proximal humerus]. *Unfallchirurg.* Juni 1999;102(6):422–8.
31. Köstler W, Strohm PC, Südkamp NP. [Proximal humerus fracture in advanced age. Treatment with fixed angle plate osteosynthesis]. *Chirurg.* November 2003;74(11):985–9.
32. Riem S, Büchele G, Becker C, Rapp K. Heimunterbringung und Mortalität nach Frakturen für zuhauselebende Senioren. In *German Medical Science GMS Publishing House*; 2016. S. DocPO02-51.
33. Rapp K, Becker C, Cameron ID, Klenk J, Kleiner A, Bleibler F, u. a. Femoral fracture rates in people with and without disability. *Age Ageing.* September 2012;41(5):653–8.
34. © Statistisches Bundesamt. Todesursachenstatistik [Internet]. 2018 [zitiert 18. Juni 2020]. Verfügbar unter:
http://www.gbe-bund.de/oowa921-install/servlet/oowa/aw92/dboowasys921.xwdevkit/xwd_init?gbe.isgbetol/xs_start_neu/&p_aid=i&p_aid=45699134&nummer=6&p_sprache=D&p_indsp=-&p_aid=13339407
35. Biber R, Grüninger S, Bail HJ. Frakturen der oberen Extremitäten im Alter. *Z Gerontol Geriat.* 2017;(50):83–5.
36. Krischak G. Traumatologie für Physiotherapeuten. 2., Auflage. Stuttgart: Thieme; 2009.
37. Tittel K. Beschreibende und funktionelle Anatomie des Menschen by Kurt Tittel. Urban & Fischer Verlag/Elsevier GmbH;

38. Menck J, Döbler A, Döhler JR. [Vascularization of the humerus]. *Langenbecks Arch Chir.* 1997;382(3):123–7.
39. Zalpour C. Für die Physiotherapie: Anatomie Physiologie. München: Urban & Fischer Verlag/Elsevier GmbH; 2002. 636 S.
40. Tillmann BN. Atlas der Anatomie des Menschen: mit Muskeltabellen [Internet]. 3. Aufl. Berlin Heidelberg: Springer-Verlag; 2016 [zitiert 25. Juni 2020]. (Springer-Lehrbuch). Verfügbar unter: <https://www.springer.com/de/book/9783662492871>
41. Schünke M, Schulte E, Schumacher U, Rude J. PROMETHEUS Lernatlas der Anatomie. Allgemeine Anatomie und Bewegungssystem. 1.,. Thieme; 2005. 542 S.
42. Kara H, Bayir A, Ak. A, Acar D, Akinci M, Degirmenci S. Trivial Trauma Induced Bilateral Proximal End Humerus Fracture: Two case reports. *Journal of Case Reports.* 20. Oktober 2013;3(2):366–9.
43. Plecko M. Über Zementaugmentation am proximalen Humerus. *OP-Journal.* 2017;(33):64–71.
44. Müller ME, Nazarian S, Koch P, Schatzker J. The Comprehensive Classification of Fractures of Long Bones. Berlin Heidelberg: Springer-Verlag; 1990.
45. AO Foundation. AO/OTA Fracture and Dislocation Classification - Introduction to the classification of long-bone fractures [Internet]. Switzerland; 2018 [zitiert 29. Juni 2020]. Verfügbar unter: https://classification.aomedical.org/files/download/AOOTA_Classification_2018_Classification_brochure_1807031108.pdf
46. Neer CS. Displaced proximal humeral fractures: part I. Classification and evaluation. 1970. *Clin Orthop Relat Res.* Januar 2006;442:77–82.
47. Voigt C, Lill H. Winkelstabile Plattenosteosynthese proximaler Humerus-frakturen. *Trauma und Berufskrankheit.* 1. März 2005;7(1):S10–4.
48. Lill H, Josten C. [Conservative or operative treatment of humeral head fractures in the elderly?]. *Chirurg.* November 2001;72(11):1224–34.
49. Lanting B, MacDermid J, Drosdowech D, Faber KJ. Proximal humeral fractures: a systematic review of treatment modalities. *J Shoulder Elbow Surg.* Februar 2008;17(1):42–54.
50. Neer CS. Displaced proximal humeral fractures. II. Treatment of three-part and four-part displacement. *J Bone Joint Surg Am.* September 1970;52(6):1090–103.
51. Tingart M, Bächli H, Bouillon B, Tilling T. [The displaced proximal humeral fracture: is there evidence for therapeutic concepts?]. *Chirurg.* November 2001;72(11):1284–91.
52. Mäurer J, Rudolph J, Jerosch J. Traumatologie. Frakturen - Frakturlokalisation. In: Mäurer J, Herausgeber. Effiziente Schulterbildgebung - Ein Update für den klinischen Alltag. Stuttgart: Georg Thieme Verlag; 2002. S. 25.

53. Habermeyer P. Die Humeruskopffraktur. Unfallchirurg. 1. Oktober 1997;100(10):820–37.
54. Lamb SE, Jørstad-Stein EC, Hauer K, Becker C, Prevention of Falls Network Europe and Outcomes Consensus Group. Development of a common outcome data set for fall injury prevention trials: the Prevention of Falls Network Europe consensus. J Am Geriatr Soc. September 2005;53(9):1618–22.
55. Orces CH. Prevalence and Determinants of Fall-Related Injuries among Older Adults in Ecuador [Internet]. Bd. 2014, Current Gerontology and Geriatrics Research. Hindawi; 2014 [zitiert 7. April 2020]. S. e863473. Verfügbar unter: <https://www.hindawi.com/journals/cggr/2014/863473/>
56. Wolke J, Scheibel M. Konservative Therapie: Wer und wie? In: Die proximale Humerusfraktur. Berlin Heidelberg: Springer-Verlag; 2014. S. 29–36.
57. Deutsche Gesellschaft für Orthopädie und Unfallchirurgie e.V. (Hrg.). Nachbehandlungsempfehlungen. Berlin: Arbeitskreis Traumarehabilitation, Sektion Physikalische Therapie und Rehabilitation; 2016.
58. Powers JH 3rd, Patrick DL, Walton MK, Marquis P, Cano S, Hobart J, u. a. Clinician-Reported Outcome Assessments of Treatment Benefit: Report of the ISPOR Clinical Outcome Assessment Emerging Good Practices Task Force. Value Health. 2017/01/10. Januar 2017;20(1):2–14.
59. Faller H, Lang H. Medizinische Psychologie und Soziologie. 3. Aufl. Berlin Heidelberg: Springer-Verlag; 2010. (Springer-Lehrbuch).
60. Döring N, Bortz J. Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften [Internet]. 5. Aufl. Heidelberg: Springer-Verlag; 2015 [zitiert 8. April 2020]. Verfügbar unter: <https://www.beck-shop.de/doering-bortz-forschungsmethoden-evaluation-sozial-humanwissenschaften/product/13383891>
61. Cronbach LJ, Warrington WG. Time-limit tests: estimating their reliability and degree of speeding. Psychometrika. Juni 1951;16(2):167–88.
62. Beaton DE. Understanding the relevance of measured change through studies of responsiveness. Spine. 15. Dezember 2000;25(24):3192–9.
63. Liang MH. Longitudinal construct validity: establishment of clinical meaning in patient evaluative instruments. Med Care. September 2000;38(9 Suppl):II84-90.
64. Liang MH, Lew RA, Stucki G, Fortin PR, Daltroy L. Measuring clinically important changes with patient-oriented questionnaires. Med Care. April 2002;40(4 Suppl):II45-51.
65. Berg K, Norman KE. Functional Assessment of Balance and Gait. Clinics in Geriatric Medicine. 1. November 1996;12(4):705–23.
66. Kazis L, Anderson J, Meenan R. Effect Sizes for Interpreting Changes in Health Status. Medical Care [Internet]. 1. März 1989 [zitiert 19. Februar 2019];27(3). Verfügbar unter: insights.ovid.com

67. Liang MH, Fossel AH, Larson MG. Comparisons of Five Health Status Instruments for Orthopedic Evaluation. *Medical Care*. Juli 1990;28(7):632.
68. Eton DT, Yost KJ. Combining Distribution- and Anchor-Based Approaches to Determine Minimally Important Differences: The FACIT Experience. *Eval Health Prof*. 1. Juni 2005;28(2):172–91.
69. Cella D, Eton DT, Fairclough DL, Bonomi P, Heyes AE, Silberman C, u. a. What is a clinically meaningful change on the Functional Assessment of Cancer Therapy-Lung (FACT-L) questionnaire?: Results from Eastern Cooperative Oncology Group (ECOG) study 5592. *Journal of clinical epidemiology*. 14. März 2002;55(3):285–95.
70. Roach KE, Budiman-Mak E, Songsiridej N, Lertratanakul Y. Development of a shoulder pain and disability index. *Arthritis Care Res*. Dezember 1991;4(4):143–9.
71. Deutsche Vereinigung für Schulter- und, Ellenbogenchirurgie (DVSE) e.V. (Hrg.). *Untersuchungstechniken des Schultergelenks. Obere Extremität*. 2010;7:3-68.
72. Magosch P, Lichtenberg S, Loew M, Tauber M, Habermeyer P. Die Klinische Untersuchung der Schulter. *Dtsch Z Sportmed [Internet]*. 1. Dezember 2013 [zitiert 8. April 2020];2013(12). Verfügbar unter: <http://www.zeitschrift-sportmedizin.de/artikel-online/archiv-2013/heft-12/die-klinische-untersuchung-der-schulter/>
73. Katzman R, Brown T, Fuld P, Peck A, Schechter R, Schimmel H. Validation of a short Orientation-Memory-Concentration Test of cognitive impairment. *Am J Psychiatry*. Juni 1983;140(6):734–9.
74. Hayes MHS, Patterson DG. Experimental development of the graphic rating method. *Psychological Bulletin*. 1921;(18):98–9.
75. Oberle C, Becker C, Schölch S, Lenz J-U, Studier-Fischer S, Augat P, u. a. Inter-rater and intra-rater reliability of an adapted Wolf motor function test for older patients with shoulder injuries. *Z Gerontol Geriat*. 1. April 2018;51(3):293–300.
76. Zech A, Hübscher M, Vogt L, Banzer W, Hänsel F, Pfeifer K. Neuromuscular training for rehabilitation of sports injuries: a systematic review. *Med Sci Sports Exerc*. Oktober 2009;41(10):1831–41.
77. Nerz C, Schwickert L, Becker C, Studier-Fischer S, Müßig JA, Augat P. Effectiveness of robot-assisted training added to conventional rehabilitation in patients with humeral fracture early after surgical treatment: protocol of a randomised, controlled, multicentre trial. *Trials*. 6. Dezember 2017;18(1):589.
78. Pereira ND, Michaelsen SM, Menezes IS, Ovando AC, Lima RCM, Teixeira-Salmela LF. Reliability of the brazilian version of the Wolf Motor Function Test in adults with hemiparesis. *Brazilian Journal of Physical Therapy*. Juni 2011;15(3):257–65.

79. Nijland R, van Wegen E, Verbunt J, van Wijk R, van Kordelaar J, Kwakkel G. A comparison of two validated tests for upper limb function after stroke: The Wolf Motor Function Test and the Action Research Arm Test. *J Rehabil Med.* Juli 2010;42(7):694–6.
80. Morris DM, Uswatte G, Crago JE, Cook EW, Taub E. The reliability of the wolf motor function test for assessing upper extremity function after stroke. *Arch Phys Med Rehabil.* Juni 2001;82(6):750–5.
81. Bogard K, Wolf S, Zhang Q, Thompson P, Morris D, Nichols-Larsen D. Can the Wolf Motor Function Test be streamlined? *Neurorehabil Neural Repair.* Juni 2009;23(5):422–8.
82. Wu C, Fu T, Lin K, Feng C, Hsieh K, Yu H, u. a. Assessing the Streamlined Wolf Motor Function Test as an Outcome Measure for Stroke Rehabilitation. *Neurorehabil Neural Repair.* 1. Februar 2011;25(2):194–9.
83. Evans JD. *Straightforward statistics for the behavioral sciences.* Pacific Grove: Brooks/Cole Pub. Co.; 1996.
84. Van Ancum J, Van Schooten KS, Jonkman NH, Huijben B, Van Lummel RC, Meskers CGM, u. a. Gait speed assessed by a 4-m walk test is not representative of daily-life gait speed in community-dwelling adults. *Maturitas.* 2019;(121):28–34.
85. Giannouli E, Bock OL, Zijlstra W. *Mobility in Old Age: Capacity Is Not Performance.* *BioMed Research International.* 2016;2016.
86. Bahadori S, Davenport P, Immins T, Wainwright TW. Validation of joint angle measurements: comparison of a novel low-cost marker-less system with an industry standard marker-based system. *J Med Eng Technol.* 2019;43(1):19–24.
87. Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: a critical review and recommendations. *Journal of Clinical Epidemiology.* 1. Mai 2000;53(5):459–68.
88. Sawilowsky S. *New Effect Size Rules of Thumb.* *Theoretical and Behavioral Foundations of Education Faculty Publications [Internet].* 1. November 2009; Verfügbar unter: https://digitalcommons.wayne.edu/coe_tbf/4
89. MacDermid JC, Richards RS, Roth JH. Distal radius fracture: A prospective outcome study of 275 patients. *Journal of Hand Therapy.* 1. April 2001;14(2):154–69.
90. Nerz C, Schwickert L, Schölch S, Gordt K, Nolte P-C, Kröger I, u. a. Inter-rater reliability, sensitivity to change and responsiveness of the orthopaedic Wolf-Motor-Function-Test as functional capacity measure before and after rehabilitation in patients with proximal humeral fractures. *BMC Musculoskeletal Disorders.* 2019;(20).

11 Erklärung zum Eigenanteil der Dissertationsschrift

Die Arbeiten im Rahmen dieser Dissertation wurde im Robert-Bosch-Krankenhaus in der Klinik für geriatrische Rehabilitation in Stuttgart unter Betreuung von Herrn Prof. Dr. Becker durchgeführt.

Die Konzeption der Studie zur Entwicklung des WMFT-O (Manuskript 1) erfolgte in Zusammenarbeit mit Herrn Prof. Dr. Becker, Chefarzt der Klinik für geriatrische Rehabilitation am Robert-Bosch-Krankenhaus in Stuttgart und Leiter der Forschungsabteilung sowie Herrn Dr. Schwickert, wissenschaftlicher Mitarbeiter der Forschungsabteilung des Robert-Bosch-Krankenhauses.

Die Konzeption der multizentrisch durchgeführten, randomisierten und kontrollierten Studie erfolgte in Zusammenarbeit mit Herrn Prof. Dr. Augat, Leiter des Instituts für Biomechanik an der Berufsgenossenschaftlichen Unfallklinik Murnau und Herrn Dr. Studier-Fischer, Leiter der Sektion für Schulter und Ellenbogenchirurgie an der Berufsgenossenschaftlichen Unfallklinik Ludwigshafen.

Die Untersuchungen der Patienten wurden in Stuttgart von mir in Zusammenarbeit mit Herrn Dr. Schwickert, Frau Schölch und Frau Dr. Gordt durchgeführt. Die Untersuchungen im Studienzentrum in Murnau wurden nach Einarbeitung von Frau Kröger und Frau Müssig unter der Leitung von Herrn Prof. Dr. Augat durchgeführt. In Ludwigshafen wurden die Messungen nach Einarbeitung von Frau Kleebach und Herrn Freischmidt unter der Leitung von Herrn Dr. Studier-Fischer und Herrn Dr. Nolte durchgeführt.

Die statistische Auswertung erfolgte nach Rücksprache eigenständig durch mich.

Ich versichere, die Dissertation selbständig verfasst zu haben und keine weiteren als die von mir angegebenen Quellen und Hilfsmittel verwendet zu haben.

Stuttgart, den

12 Weitere Publikationen und Kongressbeiträge

Publikationen

Schwickert L, **Oberle C**, Becker C, Lindemann U, Klenk J, Schwenk M, Bourke A, Zijlstra W: *Model development to study strategies of younger and older adults getting up from the floor*. Aging Clinical and Experimental Research. 2015 28(2). DOI 10.1007/s40520-015-0397-1

Gordt K, Mikolaizak S, **Nerz C**, Barz C, Gerhardy T, Weber M, Becker C, Schwenk M. *German version of the Community Balance and Mobility Scale: Translation and evaluation of measurement properties*. Z Gerontol Geriat 2018 52(1):28-36. DOI: 10.1007/s00391-018-1374-z

Weber M, Van Ancum J, Bergquist R, Taraldsen K, Gordt K, Mikolaizak S, **Nerz C**, Pijnappels M, Jonkman N, Maier A, Helbostad JL, Vereijken B, Becker C, Schwenk M. *Concurrent validity and reliability of the Community Balance and Mobility scale in young-older adults*. BMC Geriatrics 2018 18:156. DOI: 10.1186/s12877-018-0845-9

Jansen C, **Nerz C**, Kramer F, Labudek S, Klenk J, Dams J, König HH, Clemson L, Becker C, Schwenk M. *Comparison of a Group-delivered and Individually Delivered Lifestyle-integrated Functional Exercise (LiFE) Program in Older Persons: A Randomized Noninferiority Trial*. BMC Geriatrics 2018 18(1):267. DOI: 10.1186/s12877-018-0953-6

Clemson L, Munro J, Fiatarone Singh M, Schwenk M, **Nerz C**, 2018. *Trainer-Manual - Aktiv und sicher durchs Leben mit dem LiFE Programm: Evidenzbasierte Sturzprophylaxe für Senioren*. Springer Heidelberg; Auflage: 1. Aufl. 2019. ISBN-10: 3662575817

Schwenk M, Bergquist R, Boulton E, Van Ancum J, **Nerz C**, Weber M, Barz C, Jonkman N, Taraldsen K, Helbostad J, Vereijken B, Pijnappels M, Maier A, Zhang W, Becker C, Todd C, Clemson L, Hawley-Hague H. *A Programme for Preventing Functional Decline in Young Seniors: Development and Initial Evaluation*. The Gerontologist 2018. DOI: 10.1159/000499962

Gordt K, Mikolaizak S, Taraldsen K, Bergquist R, Van Ancum J, **Nerz C**, Pijnappels M, Maier A, Helbostad J, Vereijken B, Becker C, Schwenk M. *Creating and validating a shortened version of the Community Balance & Mobility Scale for the application in young seniors. Physical Therapy* 2019. <https://doi.org/10.1093/ptj/pzz132>

Kramer F, Labudek S, Jansen CJ, **Nerz C**, Fleig L, Clemson L, Becker C, Schwenk M. Development of a conceptual framework for a group-based format of the Lifestyle integrated Functional Exercise (gLiFE) programme and its initial feasibility testing. *Pilot Feasibility Stud* 6, 6 (2020). <https://doi.org/10.1186/s40814-019-0539-x>

Gordt K, Paraschiv-Ionescu A, Mikolaizak S, Taraldsen K, Mellone S, Bergquist R, Van Ancum J, **Nerz C**, Pijnappels M, Maier A, Helbostad J, Vereijken B, Becker C, Aminian K, Schwenk M. *The association of basic and challenging motor capacity with mobility performance and falls in young seniors. Archives of Gerontology and Geriatrics* 2020. <https://doi.org/10.1016/j.archger.2020.104134>

Persönliche Kongressbeiträge

Oberle C, Becker C, Schölch S, Lenz U, Studier-Fischer S, Augat P, Schwickert L. *Inter- und Intra-Reliabilität eines angepassten Wolf-Motor-Funktionstest für ältere Patienten nach Schulterverletzungen*. 28. Deutscher Geriatrie Kongress, DGGG, 08. September 2016, Stuttgart (Poster; nach Call for Papers).

Nerz C, Becker C, Studier-Fischer S, Müßig J A, Augat P, Schwickert L. *Wirksamkeit eines frühzeitigen, assistiven Robotiktrainings bei Patienten nach Humerusfrakturen: Studienprotokoll einer randomisiert kontrollierten multi-zentrischen Studie*. 29. Deutscher Geriatrie Kongress, DGG, 29. September 2017, Frankfurt (Poster; nach Call for Papers).

Nerz C, Barz C, Kroog A, Mikolaizak S, Schwenk M, Becker C. "LiFE goes on": Neue Entwicklungen und Studien zum Lifestyle-integrated Functional Exercise (LiFE) Konzept. - Nutzung von mHealth-Technologie zur Unterstützung einer alltagsintegrierten Trainingsintervention. 30. Deutscher Geriatrie Kongress, DGGG, 07. September 2018, Köln (Symposiums Beitrag; nach Call for Papers).

Nerz C, Augat P, Becker C, Schölch S, Schwickert L. *Videobasierte Überprüfung der Inter-Rater Reliabilität des angepassten Wolf-Motor-Funktionstests für orthopädische Patienten nach Schulterverletzungen*. 30. Deutscher Geriatrie Kongress, DGGG, 07. September 2018, Köln (Poster; nach Call for Papers).

Nerz C, Schwickert L, Augat P, Gordt K, Schölch S, Becker C. Sensitivity to change of the orthopaedic Wolf-Motor-Function-Test as functional capacity measure before and after rehabilitation in patients with proximal humeral fractures. MobEx, Trondheim, 18.-19. Januar 2019 (Vortrag; nach Call for Papers).

Nerz C, Schwickert L, Schölch S, Kröger I, Augat P, Becker C. Aktuelle Studien aus der Rehabilitation: Testentwicklung und Ergebnisse des WMFT-O als Schulter-Assessment in Studien und der Rehabilitation. 2. Rehasymposium der BG-Kliniken, Ludwigshafen, 26.-27. September 2019 (Workshop).

13 Danksagung

Ich möchte mich zunächst ganz herzlich bei meinem Doktorvater Herrn Prof. Dr. Clemens Becker für die Möglichkeit der Bearbeitung dieses interessanten Themas und die besondere Betreuung bedanken. Vor allem bin ich Ihm für die hervorragende Unterstützung in allen organisatorischen und inhaltlichen Belangen meiner Promotion dankbar. Er war es, der schon zu Beginn meines Praktikums am Robert-Bosch-Krankenhaus an mich geglaubt hat und mir durch seine berufliche Förderung erst die Promotion ermöglicht hat.

Zu ganz besonderem Dank bin ich Herrn Dr. Lars Schwickert verpflichtet, der mich während meiner gesamten Promotion stets in allen Belangen unterstützt und motiviert hat und mir immer mit Rat und Tat zur Seite stand. Danke für die tolle Zusammenarbeit und die daraus entstandene Freundschaft.

Die Durchführung der Studie wäre ohne die Hilfe von Frau Sabine Schölch und Frau Dr. Katharina Gordt, sowie den Kollegen und Kolleginnen aus den berufsgenossenschaftlichen Kliniken in Murnau und Ludwigshafen nicht möglich gewesen. Für Ihre Unterstützung bei der Durchführung der Assessments und der Beurteilung der Probanden in den klinischen Ratings möchte ich mich bedanken. Ein großes Dankeschön geht auch an Frau Eva Ermer, Frau Rebekka Leonhardt und Herrn Christoph Endress, welche mir bei der Beurteilung der Probanden des videobasierten Wolf-Motor-Funktionstests unter die Arme gegriffen haben.

Mein Dank geht an alle Probanden, welche sich bereiterklärt haben an einer Studie im Rahmen meiner Promotion teilzunehmen. Ohne sie wäre meine Promotion nicht möglich gewesen.

Des Weiteren möchte ich die Möglichkeit nutzen und mich bei meinen Kolleginnen und Kollegen der Forschungsabteilung des Robert-Bosch-Krankenhauses für die wertvollen Anregungen, die konstruktive Kritik und die fachliche wie auch menschliche Unterstützung im Laufe meiner Promotion bedanken. Ihr seid ein tolles Team, in welchem es Spaß macht täglich zu arbeiten.

Ein besonderer Dank geht hier an Kim Sczuka, Dr. Katharina Gordt, Dr. Stefanie Mikolaizak, Hella Kircheisen und Christoph Endress, welche mir moralisch zur Seite

standen, auch wenn es schwierig wurde und mich immer unterstützt und an mich geglaubt haben. Ich freue mich euch nicht nur zu meinen Kollegen, sondern auch zu meinen Freunden zählen zu dürfen. Dank eurer positiven und optimistischen Art und euren vielen lieben Worte während meiner Promotion habt Ihr mir geholfen, alle Herausforderungen zu meistern und daran zu wachsen.

Von Herzen möchte ich mich bei meinen Freundinnen Katharina Hein, Juliane Spachmann und Laura Czechanowski bedanken. Sie hatten immer ein offenes Ohr, waren immer für mich da und haben mich mit vielen lieben Worten in allem unterstützt.

Zu guter Letzt möchte ich mich von ganzem Herzen bei meiner Familie bedanken, die mir immer geholfen und mich immer unterstützt hat. Insbesondere bei meinem Vater Roland Oberle der alle meine Arbeiten Korrektur gelesen hat. Es freut und ehrt mich, dass ihr meinen beruflichen Werdegang mit so viel Stolz verfolgt.

Ein ganz besonderer Dank geht an meinen Ehemann Benjamin Nerz, der mich immer ermutigt hat, immer an mich geglaubt hat, immer hinter mir stand und mir in jeder Situation Halt gegeben hat. Er war in herausfordernden und schwierigen Situationen für mich da, hat mir Rückhalt gegeben und hat sich über jeden auch noch so kleinen Fortschritt und Erfolg mit mir gefreut!

Vielen Dank für alles!

14 Anhang: Manuskripte 1-4

Manuskript 1:

Oberle C, Becker C, Schölch S, Lenz JU, Studier-Fischer S, Augat P, Schwickert L. Inter-rater and intra-rater reliability of an adapted Wolf motor function test for older patients with shoulder injuries. *Z Gerontol Geriat.* 2017. DOI 10.1007/s00391-017-1226-2

Nachdruck mit Genehmigung (4867600636332) von "Copyright Clearance Center": Springer Nature; Zeitschrift für Gerontologie und Geriatrie; Inter-rater and intra-rater reliability of an adapted Wolf motor function test for older patients with shoulder injuries; Corinna Oberle et al. COPYRIGHT © 2017

Z Gerontol Geriat
 DOI 10.1007/s00391-017-1226-2
 Received: 3 January 2017
 Revised: 27 February 2017
 Accepted: 13 March 2017
 © Springer Medizin Verlag GmbH 2017



Corinna Oberle¹ · Clemens Becker¹ · Sabine Schölch¹ · Joerg-Uwe Lenz² · Stefan Studier-Fischer³ · Peter Augat^{4,5} · Lars Schwickert¹

¹ Department for Clinical Gerontology, Robert-Bosch-Hospital, Stuttgart, Germany

² Department for Orthopaedics and Trauma Surgery, Robert-Bosch-Hospital, Stuttgart, Germany

³ Department of Trauma and Orthopedic Surgery, BG Trauma Center Ludwigshafen, Ludwigshafen, Germany

⁴ Institute of Biomechanics, BG Trauma Center Murnau, Murnau, Germany

⁵ Institute of Biomechanics, Paracelsus Medical University, Salzburg, Austria

Inter-rater and intra-rater reliability of an adapted Wolf motor function test for older patients with shoulder injuries

Introduction

Shoulder injuries constitute a frequent health problem in older adults being associated with severe individual as well as societal economic consequences. Proximal humeral fractures (PHF) increase by number, particularly in adults 70 years and older [1, 2]. A recent study in southern Germany showed that the mean age of PHF patients presenting in hospitals was 80.1 years and PHF represented 14% of all patients hospitalized with fractures [3]. Until now, there is no evidence-based consensus on rehabilitation after PHF [4–6]. Heterogeneous rehabilitation concepts exist without standardization of content, duration, intensity and frequency. This is accompanied by a lack of validated clinical instruments which can objectively and reliably assess functional capacity (FC) and task performance of the upper extremities after shoulder injuries and PHF. Commonly used assessment instruments either use subjective ratings (SPADI [7], DASH [8, 9], SVISS [10], SST; [11], SSRS [12], ASES [13]) or partly involve objective measurements (Constant score [14]). This has been shown to be questionable in e. g. sensitivity to change and other test measurement properties [15]. Hence, there is a need for better and objective assessments of shoulder function to assess patients and mea-

sure rehabilitation progress. A promising neurological tool assessing the functional capacity and quality of movement (QoM) of the upper extremities is the Wolf motor function test (WMFT-N). Originally it was developed to assess upper extremity motor deficits due to neurological disorders and traumatic brain injuries [14]. The original test was later modified [16] to consider different severity levels of impairments. Further modifications were provided in [17] adding a scale to assess the QoM. The test measurement properties, in terms of inter-test reliability, inter-rater reliability, internal consistency and stability were deemed excellent for both the performance time and FC rating ranging from 0.88 to 0.98 [16, 18]. A pilot study performed by our group has shown the overall applicability of the neurological based WMFT-N to orthopedic patients with shoulder injuries but identified some shortcomings [19]. Replacing distal tasks, highly relevant for patients with neurological disorders by more proximal tasks was one aim. Also challenging the arm to move against gravity was further recommended. Hence, the aim of this study was to assess the inter-rater and intra-rater reliability of the WMFT-O which adapted the WMFT-N towards the assessment of the orthopedic shoulder injuries.

Methods

Adaptation process

The 17 item WMFT-N version of Taub et al. as mentioned in [16] served as a basis for development of the WMFT-O. The WMFT-O included 18 upper extremity motor tasks with increased focus on proximal tasks involving the shoulder joint compared to the dexterity tasks addressing finger and hand movement (Table 1). In total, three tasks involving abduction in the shoulder joint (task 1, 2ab, 3), one task for abduction and external rotation in the shoulder joint (task 18), two tasks for extension of the elbow joint (task 4, 5) and one task for flexion at the elbow joint (task 9) as well as three tasks addressing complex movements in the shoulder joint (task 6, 7ab, 8), one task in which the grip strength was measured (task 15) and seven tasks involving activities of daily living were evaluated (task 10–14, 16, 17; <https://www.youtube.com/watch?v=8yDcWK9Xrtw>; <https://www.youtube.com/watch?v=MsfGauYHfc>).

If the patient reached the 30 cm high box without extra weight in tasks 2 and 7, the following tasks (3 and 8) including extra weight were performed with a 30 cm high box, otherwise the task was

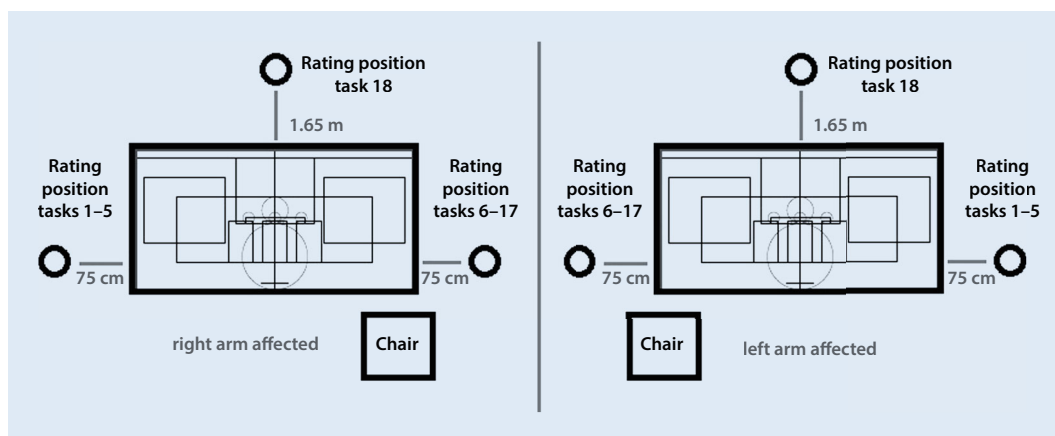


Fig. 1 ◀ Rating positions for each task and side of the affected arm. Patients are seated on the chair

performed with a 15 cm high box. If the subject previously was not able to place the arm on the 15 cm high box without weight, the task including extra weight was skipped.

Maximum grip strength of both hands was measured three times using a handheld Jamar dynamometer as suggested in the literature [16, 20–22].

For each performed task, functional capacity (FC) and quality of movement (QoM) were rated according to a 6-point Likert scale. For developing a rating scale for grip strength, categories were created according to average values from a meta-analysis of grip strength [23]. For each age group (65–69, 70–74 and 75+ years) the respective average served as the upper limit of the rating scale. Based on this reference values 6 intervals (from “0” to the respective average) were graduated to achieve a 6-point Likert scale. The required time to complete the individual tasks was taken by one rater with a stopwatch. In agreement with Whittall et al. the maximum time to perform a task was 120 s and the task was stopped when the subject passed this threshold [24]. During all tasks using a lateral sitting position in relation to the table, the raters stood laterally to the table on the non-affected arm side of the subject so that the rater and the subject stood face to face. During tasks, which were carried out in a frontal sitting position in relation to the table, the raters stood laterally to the table on the affected arm side of the subject to enable a non-occluded view of the affected shoulder. The only exception was the last task [18] with the raters standing

frontal to the table at a distance of 1.65 m looking towards the subject (◻ Fig. 1).

Standardized instructions were read to the subjects once before each task. In addition, each task was demonstrated and explained by the investigator once. The subjects were not allowed to practice the movement. If the subjects confirmed to understand, the tasks and time measurement were initiated with the signal words “on your mark” – “ready” and “go”. The end point was defined by the completion of the predefined tasks.

Assessment procedure

For testing inter-rater reliability, a sample of patients to geriatric rehabilitation was assessed by two out of three trained raters. For intra-rater reliability, the same sample was assessed twice by the same rater, measurements were applied within 2 days to minimize the effect of spontaneous recovery. Exclusion criteria were cognitive impairment, short orientation memory concentration test (SOMC) test ≥ 10 , strong visual impairment (vision < 0.29) and pain during implementation of the WMFT-O or injury of the axillary nerve. All participants gave written informed consent. The study was approved by the ethical committee of the University of Tübingen and is in agreement with the Declaration of Helsinki.

Statistics

Both, inter-rater and intra-rater agreement were determined by weighted Cohen’s Kappa (K_w) statistics with corresponding confidence intervals (CI) as

suggested for ordinal scaled data (underlying two-way-random effect models with absolute agreement definition) [25]. The WMFT-O was considered to have consistent factor loading on the 18 items. Cronbach’s alpha was calculated to determine the internal consistency between the items of each scale. All Cohen’s kappa statistics were calculated with SAS® Studio (University edition 3.4, Cary, NC, USA). Descriptive statistics and Cronbach’s alphas were calculated with SPSS (Version 23.0).

Results

The study sample consisted of 20 older subjects (16 women, 4 men) with a median age of 80.5 years (range 64–94 years). There were no adverse events or complications in this study. The subject characteristics are listed in ◻ Table 2. Due to the small proportion of men in this target group no gender segregation has been made for further calculations.

Inter-rater reliability

◻ Table 3 lists inter-rater reliability in terms of K_w as well as internal consistency for FC and QoM for each task. All tasks showed “good” inter-rater reliability. The K_w values ranged from 0.84–1.00 (CI = 0.64–1.00) for the FC and from 0.79–1.00 (CI = 0.57–1.00) for the QoM. The internal consistency was considered “excellent”. The Cronbach’s alpha values varied between 0.94–1.00 with a median of 0.98 for the FC and between 0.96–1.00 with a median of 0.99 for the QoM. Task 15 was omitted

because only one rater measured the grip strength, thus it could not contribute to the analysis of reliability of the test.

Intra-rater reliability

■ **Table 3** lists intra-rater reliability. The K_w values varied between 0.71 and 1.00 (CI = 0.47 to 1.00) for the FC and between 0.71 and 1.00 (CI = 0.44 to 1.00) for the QoM. Due to strong variations of the grip strength values of the affected and non-affected side (task 15) from the first to the second measurement, no K_w values were calculated. Due to non-normal distribution, intra-rater reliability was not calculated for performance time.

Discussion

This study adapted the original WMFT-N to a WMFT-O version for individuals with shoulder injuries finding good intra-rater and inter-rater reliability and internal consistency indicated by K_w analysis.

Inter-rater reliability

The calculated K_w values for the FC and QoM items ranged from 0.84–1.00 and from 0.79–1.00, respectively. According to Viera and Garrett, Cohen's kappa values of 1.00 indicate "perfect" concordance, while kappa values >0.60 indicate substantial concordance and kappa values >0.80 an "almost perfect" concordance and thus "good" inter-rater reliability [26]. These results were in line with existing studies investigating into reliability of neurological based WMFT-N versions. The Intra-Class-Correlation (ICC) values for the FC for the Brazilian version of WMFT-N showed "good" to "very good" inter-rater reliability [27] and Nijland et al. found very good reliability of the FC and the QoM of the WMFT-N in a sample with 40 patients after stroke [28]. In general, reliability can thus be assumed towards potential application as clinical score in orthopedic rehabilitation especially for patients with PHF. Single items showed lower agreement K_w values for the QoM, namely the tasks "hand to table frontal" (task 6) and "flip cards frontal" (task 14) but the rater agreement was deemed substantial ($K_w =$

Z Gerontol Geriat DOI 10.1007/s00391-017-1226-2
© Springer Medizin Verlag GmbH 2017

C. Oberle · C. Becker · S. Schölich · J.-U. Lenz · S. Studier-Fischer · P. Augat · L. Schwickert

Inter-rater and intra-rater reliability of an adapted Wolf motor function test for older patients with shoulder injuries

Abstract

Background and objective. The Wolf motor function test (WMFT) evaluates the upper extremity performance of patients with neurological disorders and traumatic brain injuries by assessing time and functional capacity. It was later modified to also measure the quality of the performed movements. This study aimed to adapt the WMFT for patients with shoulder injuries and to evaluate the inter-rater and intra-rater reliability of this adapted version.

Subjects and methods. A total of 20 individuals with a median age of 80.5 years were assessed using a test-retest design after surgical or conservative treatment of shoulder injuries. Two raters rated performance to determine inter-rater reliability. One rater rated the performance again to determine intra-rater reliability. Both inter-rater and intra-rater agreement were determined

by weighted Cohen's kappa statistics with corresponding confidence intervals. Cronbach's alpha was calculated to determine internal consistency.

Results. The inter-rater Cohen's kappa values ranged from 0.84–1.00 for functional capacity and from 0.79–1.00 for quality of movement. For intra-rater reliability the Cohen's kappa ranged from 0.71–1.00. Cronbach's alpha was >0.94 for functional capacity and >0.96 for quality of movement.

Conclusion. The adapted version of the WMFT shows high intra-rater and inter-rater reliability for patients with shoulder injuries.

Keywords

Wolf motor function test · Reliability · Orthopedic assessment · Shoulder function · Proximal humeral fracture

Testgütekriterien des modifizierten Wolf-Motor-Funktionstests für ältere Patienten nach Schulterverletzung

Zusammenfassung

Hintergrund und Zielsetzung. Der Wolf-Motor-Funktionstest (WNFT) bewertet objektiv die Leistungsfähigkeit der oberen Extremitäten von Patienten mit neurologischen Störungen und Schädel-Hirn-Verletzungen anhand der benötigten Zeit zur Ausführung bestimmter Aufgaben und der funktionellen Fähigkeit. Nach einer durchgeführten Modifikation wird die Bewegungsqualität ebenfalls bewertet. Ziel dieser Studie ist es, den WMFT auf Patienten nach Schulterverletzungen anzupassen und die Inter- und Intrarater-Reliabilität dieser angepassten Version zu überprüfen.

Patienten und Methodik. Bei 20 älteren Patienten (Median: 80,5 Jahre) mit chirurgisch oder konservativ behandelter Schulterverletzung wurde die Leistungsfähigkeit der Schulter in einem Test-Retest-Design erhoben. Um die Interrater-Reliabilität zu bestimmen bewerteten zwei unabhängige Rater die Leistung. Einer der Rater bewertete die Leistung erneut, um die Intrarater-

Reliabilität zu erheben. Inter- und Intrarater-Reliabilität wurden durch gewichtete Cohens-Kappa-Statistiken mit entsprechenden Konfidenzintervallen überprüft. Zur Bestimmung der internen Konsistenz wurde Cronbachs Alpha berechnet.

Ergebnisse. Für die Interrater-Reliabilität reichen die Cohens-Kappa-Werte für die funktionelle Fähigkeit von 0,84–1,00 und für die Bewegungsqualität von 0,79–1,00. Für die Intrarater-Reliabilität liegen die Werte bei 0,71–1,00. Die Cronbachs-Alpha-Werte lagen für die funktionelle Fähigkeit über 0,94 und für die Bewegungsqualität über 0,96.

Schlussfolgerung. Die angepasste Version des WMFT zeigt eine gute Intra- und Interrater-Reliabilität für Patienten nach Schulterverletzungen.

Schlüsselwörter

Wolf Motor Function Test · Reliabilität · Orthopädische Assessments · Schulterfunktion · Proximale Humerusfraktur

Table 1 Conducted adaptations of the WMFT-N to the WMFT-O. Adaptations written in italic

Tasks of the WMFT-N		Adapted tasks of the WMFT-O	Notes & reason for adaptation
1) Forearm to table laterally		1) Forearm to table laterally	–
2) Forearm to box 30 cm laterally	Adaption →	2a) Forearm to box 30 cm laterally	–
		2b) Forearm to box 15 cm laterally	To enable patients with limited capacity to abduct the arm to complete the task with a 15 cm high box in case they could not reach the predetermined 30 cm
	Addition →	3) Forearm to box laterally with weight 15/30 cm	Analogous to the original task 7, the applied weight was increased stepwise (in 100 g steps) starting at 100 g
3) Extend elbow laterally		4) Extend elbow laterally	–
4) Extend elbow laterally with weight		5) Extend elbow laterally with weight	–
5) Hand to table frontal		6) Hand to table frontal	–
6) Hand to box 30 cm frontal		7a) Hand to box 30 cm frontal	–
	Adaption →	7b) Hand to box 15 cm frontal	To enable patients with limited capacity to abduct the arm to complete the task with 15 cm high box in case they could not reach the predetermined 30 cm
7) Hand to box frontal with weight 30 cm		8) Hand to box frontal with weight 30 cm	The applied weight was increased stepwise (in 100 g steps) starting at 100 g
	Adaption →	Hand to box frontal with weight 15 cm	To enable patients with limited capacity to abduct the arm to complete the task with 15 cm high box in case they could not reach the predetermined 30 cm
8) Reach and retrieve frontal	Adaption →	9) Reach and retrieve frontal	Retrieve the sand bag by putting their hand flat on it and flex the elbow without performing wiping motions
9) Lift can frontal		10) Lift can frontal	–
10) Lift pencil frontal		11) Lift pencil frontal	–
11) Lift paper clip frontal		12) Lift paper clip frontal	–
12) Stack checkers frontal	Adaption →	13) Stack checkers frontal	Conducted with 6 cm high logs of wood (total stack height of 18 cm compared to 1.5–2.4 cm in the WMFT-N)
13) Flip cards frontal		14) Flip cards frontal	–
14) Grip strength affected & not affected side	Adaption →	15) Grip strength affected & not affected side	Seating position according to the standard procedure for the measurement of grip strength
15) Turn key in lock frontal		16) Turn key in lock frontal	–
16) Fold towel frontal		17) Fold towel frontal	–
17) Lift basket frontal		18) Lift basket frontal	–

0.79) and the limit to “almost perfect” was just missed on the basis of the weighted calculation of values without limiting the overall reliability of the instrument by its clinical applicability. The compliance of the ratings of the FC for the single tasks was high and in line with the original version [16]. Morris et al. [16] also reported nine tasks without adequate inter-rater reliability ($ICC_{[3,1]} < 0.75$) different to our study, thus pointing towards better applicability for the complete assessment including all tasks. The purpose-specific modifications and objectifications of the rating scales, which have been defined in detail, were considered as one potential

cause for the high inter-rater reliability in this study. Another cause may have been the adjustment of the tasks themselves. Three of the nine tasks which according to Morris et al. [16] had no adequate reliability, have been adapted for the WMFT-O.

According to George a Cronbach’s alpha between 0.80 and 0.90 was considered good, and >0.90 was considered excellent [29]. The Cronbach’s alpha values presented in this study showed an excellent internal consistency for both subscales (FC: $\alpha = 0.94$ –1.00; QoM: $\alpha = 0.96$ –1.00) and therefore exhibit a high scale reliability. Furthermore, this indicated that the

individual tasks of the WMFT-O measured the same characteristics. Hence, it was deemed feasible to summarize the results of all individual tasks to obtain an overall result.

Intra-rater reliability

The K_w values for the FC and the QoM ranged from 0.71–1.00 and therefore showed an almost perfect agreement for all tasks of the FC except for the tasks “hand to table frontal” (task 6) and “stack checkers frontal” (task 13) showing substantial agreement ($\kappa > 0.61$). In both tasks four subjects differed by one

Table 2 Subject characteristics

ID	Sex	Age (years)	Height (cm)	Weight (kg)	BMI	Side of fracture	Handedness	Injury classification (ICD 10 codes)	Treatment	Barthel index	Vision	SOMC	DASH	Pain score (VAS, T1)	Pain score (VAS, T2)
1	F	75	155	51.0	21.2	L	R	S.42.21	Surgical	75	0.80	6	21.55	1.5	2.0
2	F	64	156	50.0	20.5	R	R	S.43.00	Surgical	85	0.29	2	68.97	2.0	1.0
3	F	77	145	63.0	30.0	R	L	S.42.21	Surgical	90	0.40	3	45.69	3.0	4.0
4	F	84	160	55.6	21.7	L	R	S.42.3	Surgical	65	0.50	4	81.03	3.5	4.0
5	F	87	145	53.0	25.2	R	R	S.42.21, S.42.3	Surgical	85	0.40	9	33.63	4.5	4.5
6	M	87	182	72.0	21.7	R	R	S.42.3	Conservative	65	0.40	4	40.52	4.0	6.0
7	F	85	170	55.3	19.1	R	R	S.42.20	Surgical	30	0.40	6	62.90	3.5	4.0
8	F	87	160	61.7	24.1	R	R	S.42.20	Surgical	70	0.29	2	16.40	1.0	1.0
9	F	81	162	87.1	33.2	L	R	S.49.8, S.46.0	Surgical	81	0.50	2	71.55	0.5	0.5
10	M	77	176	76.2	24.6	R	R	S.42.3, S.42.21	Surgical	55	0.29	6	51.73	0.5	0.5
11	F	94	159	56.9	22.5	R	R	S.42.21	Conservative	70	0.50	2	38.79	1.5	1.5
12	F	80	167	87.3	31.3	R	R	S.42.21	Surgical	50	0.67	6	49.14	5.0	5.0
13	F	90	152	62.0	26.8	R	R	S.42.20	Surgical	60	0.50	3	43.97	0.0	0.0
14	M	80	160	60.5	23.6	L	R	S.42.3	Surgical	65	0.67	9	34.48	8.0	8.0
15	F	84	163	52.8	19.9	R	R	S.42.3	Surgical	70	0.50	4	29.31	4.0	3.5
16	F	83	153	68.5	29.3	R	L	S.42.21	Surgical	75	0.50	8	49.98	0.0	0.5
17	F	75	160	55.0	21.5	R	L	S.42.21, S.42.3	Surgical	65	0.50	4	43.10	0.0	0.0
18	F	79	153	56.8	24.3	R	L	S.42.21	Surgical	90	0.50	2	38.79	0.0	0.5
19	M	79	164	64.1	23.8	R	R	S.42.21, S.43.00	Surgical	90	0.50	6	30.00	0.0	0.0
20	F	80	157	63.0	25.6	R	R	S.42.2	Surgical	60	0.40	8	87.07	5.0	5.0
Median	-	80.5	160	61.1	24.0	-	-	-	-	70	0.50	4	43.54	1.75	1.75
25% Quartile	-	78.5	155	55.23	21.7	-	-	-	-	63.75	0.40	2.75	34.27	0.38	0.5
75% Quartile	-	85.5	163	65.2	25.9	-	-	-	-	77.5	0.50	6	54.52	4.00	4.13
Min	-	64	1.45	50	19.1	-	-	-	-	30	0.29	2	16.40	0.0	0.0
Max	-	94	1.82	87.2	33.2	-	-	-	-	90	0.80	9	81.07	8.0	8.0

ID: identification, BMI: body mass index, SOMC: short orientation memory concentration test, DASH: disability of the arm, shoulder and hand, VAS: visual analog scale, T1: first measurement, T2: retest

Table 3 Inter- and intra-rater reliability across all items of the WMFT-O

Task	FC			QoM		
	Inter-rater reliability K_w (95% CI)	Intra-rater reliability K_w (95% CI)	Internal consistency α	Inter-rater reliability K_w (95% CI)	Intra-rater reliability K_w (95% CI)	Internal consistency α
1. Forearm to table lateral	0.86 (0.67–1.00)	0.93 (0.79–1.00)	0.97	0.93 (0.80–1.00)	1.00 (1.00–1.00)	0.99
2a/b. Forearm to box 15/30 cm lateral	1.00 (1.00–1.00)	0.86 (0.67–1.00)	0.99	0.86 (0.68–1.00)	0.93 (0.80–1.00)	0.99
3. Forearm to box lateral with weight	0.92 (0.78–1.00)	0.84 (0.62–1.00)	1.00	0.93 (0.80–1.00)	0.79 (0.56–1.00)	1.00
4. Extend elbow lateral	0.86 (0.70–1.00)	0.85 (0.69–1.00)	0.94	0.93 (0.80–1.00)	0.80 (0.59–1.00)	0.99
5. Extend elbow lateral with weight	0.93 (0.78–1.00)	0.85 (0.66–1.00)	0.95	1.00 (1.00–1.00)	0.74 (0.51–0.97)	1.00
6. Hand to table frontal	0.86 (0.67–1.00)	0.71 (0.47–1.00)	0.97	0.79 (0.58–1.00)	0.86 (0.68–1.00)	0.96
7a/b. Hand to box 15/30 cm frontal	1.00 (1.00–1.00)	0.86 (0.68–1.00)	1.00	0.93 (0.79–1.00)	0.72 (0.48–1.00)	0.99
8. Hand to box frontal with weight	0.93 (0.78–1.00)	0.85 (0.66–1.00)	0.99	0.93 (0.81–1.00)	0.73 (0.50–1.00)	0.99
9. Reach and retrieve frontal	0.87 (0.71–1.00)	0.82 (0.62–1.00)	0.98	0.93 (0.78–1.00)	0.77 (0.51–1.00)	0.99
10. Lift can frontal	0.92 (0.77–1.00)	0.85 (0.66–1.00)	0.99	0.86 (0.69–1.00)	0.93 (0.80–1.00)	0.98
11. Lift pencil frontal	1.00 (1.00–1.00)	0.93 (0.80–1.00)	1.00	0.93 (0.79–1.00)	0.71 (0.44–0.98)	0.99
12. Lift paper clip frontal	0.84 (0.64–1.00)	1.00 (1.00–1.00)	0.98	0.86 (0.68–1.00)	1.00 (1.00–1.00)	0.99
13. Stack checkers frontal	0.93 (0.80–1.00)	0.72 (0.49–1.00)	0.99	0.94 (0.81–1.00)	0.87 (0.70–1.00)	0.99
14. Flip cards frontal ¹	0.85 (0.65–1.00)	0.92 (0.77–1.00)	0.98	0.79 (0.57–1.00)	0.86 (0.68–1.00)	0.97
16. Turn key in lock frontal	1.00 (1.00–1.00)	0.84 (0.64–1.00)	1.00	0.87 (0.69–1.00)	0.80 (0.58–1.00)	0.98
17. Fold towel frontal	0.87 (0.70–1.00)	0.93 (0.80–1.00)	0.98	0.87 (0.71–1.00)	0.94 (0.81–1.00)	0.99
18. Lift basket frontal	0.85 (0.64–1.00)	0.93 (0.78–1.00)	0.98	0.94 (0.82–1.00)	0.87 (0.71–1.00)	0.99

FC functional capacity, QoM quality of movement, K_w weighted kappa, CI confidence Interval, internal consistency α , ¹ task 15 was omitted

point from T1 to T2. In three individuals, the higher rating could be explained by an improved performance, i.e. a change from significant compensation strategies towards slight compensation strategies or almost normal movement towards normal movement. Furthermore, learning effects and day to day variation can be explanations. Another explanation was that the subjects required only an average of 2.32 s (except subject 21 T1, did not complete the tasks) to complete the task “hand to table frontal”, which is a very short time for the raters to assess differences within the task performance.

As a result, there may be a shift in the rating by one point. In the study from Pereira et al. the kappa values of the FC ranged from 0.79–0.96 [27]. The lower limit was slightly higher, the upper limit somewhat lower than in this study but mainly comparable, showing good to excellent reliability. All tasks of the QoM except seven showed an almost perfect agreement. These seven tasks (3, 4, 5, 7ab, 8, 9, 11) of the QoM showed substantial agreement ($K_w = 0.71–0.80$). The ratings of these tasks differed from T1 to T2 between 2–5 points. In 20 cases there was a higher rating in T2 as in T1,

in 5 cases the rating in T2 was lower than in T1. Hence, increase of the performance from T1 to T2, variation in their daily mood or learning effects could be possible explanations for the substantial agreement. This would suggest higher impact of daily mood on QoM than on FC.

Experience in conducting the WMFT-O

Completion time of simple proximal tasks was much shorter in patients with shoulder injuries in this study compared

to stroke patients of previous investigations. Some patients could not perform single tasks due to poor performance status. Hence the required performance time was assumed to be 120 s for those tasks. Accordingly, the measured time was not normally distributed and no inter-rater statistics were calculated. Additionally, these very short intervals lead to a lack of sensitivity of the stopwatch measurements; therefore, completion time should not be considered as an appropriate outcome measure in patients with shoulder injuries. Educated from similar sensitivity problems in other assessments, i. e. timing raising from a chair in the chair rising test and as suggested in Schwickert et al. [19], a possible solution could be a higher number of movement repetitions. Otherwise, that could lead to learning effects and muscle fatigue (peripheral fatigue), which in turn distorts the results. Another problem could be the fatigue of the subjects throughout the study (central fatigue). Due to a total of 18 motion tasks, 2 of them with additional weight in form of a weight cuff and one with additional weight in form of a beverage can (333 g) fatigues the muscles and also decreases the awareness. Increased compensational movement behaviour and relieving postures to preserve active muscle regions could be the result, which consequently lead to a lower rating. A solution to this could be the development of a shortened version of the WMFT-O, which contains less redundant movement tasks in order to reduce the burden on the subject. An attempt was already presented in the literature. Bogard et al. showed that the original version of the WMFT-N can be shortened from 17 to 6 tasks [30]. Wu et al. [31] further showed that the short version of the WMFT-N demonstrated “good” concurrent validity in participants with subacute stroke and that the predictive validity of the data obtained with the short WMFT-N was similar to the original one. Also, the responsiveness of both tests (the short and the original one) was not significantly different. These results can be interpreted as an indicator for the usefulness of a shortened WMFT-O version for patients with shoulder

injuries which is highly recommended in fracture patients who tend to suffer stronger fatigue during the full WMFT-O which took about 40 min to conduct. A suggestion therefore will be the inclusion of lateral motion tasks.

The instruction of task 8, which is based on Taub et al. [32] needs to be adjusted so that the forward inclination of the body, which evolves from the large predetermined distance between chair and table does not adversely influence the rating (due to compensatory bending forward).

Another modification which can be suggested is that the surface of table should have a low skin adhesion, because subjects with short-sleeved tops and a poor performance status had problems with the tasks 3 and 4 for which they have to move their arm across the table. The skin of the subjects stuck to the table and thus resulted into short interruptions in the movement which led to a lower rating. One solution for this problem as already mentioned in older versions of the WMFT-N could be the use of baby powder.

In terms of cognitive load the cognitive ability of the subjects to follow the instructions and repeat motor movements remains to be estimated individually by the raters. The assumed cut off for the SOMC test in this study was at a value of ≥ 10 points, while 6 points and more suggest a slight cognitive impairment [33]. Of the subjects in this study 3 had values of 6 or more points but these subjects had no problems in understanding the instructions of the WMFT-O. Consequently, the WMFT-O could probably also be used in subjects with mild cognitive impairment.

A limitation in this study was that the reliability analysis was conducted with a low sample size. Findings should therefore be replicated within a bigger sample also including younger subjects.

Conclusion

The adapted version WMFT-O for orthopedic patients is an objective, practicable and reliable clinical assessment of shoulder function after shoulder injury including proximal humeral fractures. Further analyses will involve test-

ing of younger patients with PHF, the sensibility of change of WMFT-O, video analysis and principle component analysis.

Clinical message

- The adapted version of WMFT-O is a feasible, objective and reliable measurement tool for shoulder function for older patients with shoulder injuries.
- The WMFT-O is a valid measurement tool to assess rehabilitation progress in clinical application.
- Further studies should consider larger and younger subject groups to improve generalisability.

Corresponding address



C. Oberle

Department for Clinical Gerontology, Robert-Bosch-Hospital
Auerbachstr. 110,
70376 Stuttgart, Germany
corinna.oberle@rbk.de

Acknowledgements. The study was funded and supported by the German Statutory Insurance (Deutsche Gesetzliche Unfallversicherung, DGUV). The DGUV had no influence on data assessment or writing of the manuscript.

Compliance with ethical guidelines

Conflict of interest. C. Oberle, C. Becker, S. Schölich, J.-U. Lenz, S. Studier-Fischer, P. Augat and L. Schwickert declare that they have no competing interests.

All procedures performed in studies involving human participants were approved by the ethical committee of the University of Tübingen and are in accordance with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

References

1. Resch DH (2003) Die Humeruskopffraktur. Unfallchirurg 106(8):602–617
2. Raschke MJ (2009) Alterstraumatologie: Prophylaxe, Therapie und Rehabilitation. Elsevier, Urban & Fischer, Amsterdam, p 586
3. Riem M et al (2016) Heimunterbringung und Mortalität nach Frakturen für zuhauselebende Senioren. <http://www.egms.de/static/en/meetings/altra2016/16altra16.shtml>. Accessed 18. Apr 2016

4. Zech A et al (2009) Neuromuscular training for rehabilitation of sports injuries: a systematic review. *Med Sci Sports Exerc* 41(10):1831–1841
5. Franke DS, Ambacher T (2012) Die proximale Humerusfraktur. *Obere Extremität* 7(3):137–143
6. Handoll HH, Ollivere BJ (2010) Interventions for treating proximal humeral fractures in adults. *Cochrane Database Syst Rev*. doi:10.1002/14651858.cd000434.pub2
7. Roach KE et al (1991) Development of a shoulder pain and disability index. *Arthritis Care Res* 4(4):143–149
8. Germann G et al (2003) Standardisierung und Validierung der deutschen Version 2.0 des "Disability of Arm, Shoulder, Hand" (DASH)-Fragebogens zur Outcome-Messung an der oberen Extremität. *Unfallchirurg* 106(1):13–19
9. Dixon D et al (2008) The Disabilities of the Arm, Shoulder and Hand Questionnaire (DASH) can measure the impairment, activity limitations and participation restriction constructs from the International Classification of Functioning, Disability and Health (ICF). *BMC Musculoskeletal Disord* 9(1):114
10. Boehm D et al (2004) Entwicklung eines Fragebogens basierend auf dem Constant-Murely-Score zur Selbstevaluation der Schulterfunktion durch den Patienten. *Unfallchirurg* 107(5):397–402
11. Godfrey J (2007) Reliability, validity, and responsiveness of the simple shoulder test: psychometric properties by age and injury type. *J Shoulder Elbow Surg* 16(3):260–267
12. Kohn D, Geyer M (1997) The subjective shoulder rating system. *Arch Orthop Trauma Surg* 116(6–7):324–328
13. Richards RR et al (1987) A standardized method for the assessment of shoulder function. *J Shoulder Elbow Surg* 3(6):347–352
14. Constant CR, Murley AH (1987) A clinical method of functional assessment of the shoulder. *Clin Orthop* 214:160–164
15. Conboy VB et al (1996) An evaluation of the Constant-Murley shoulder assessment. *J Bone Joint Surg Br* 78(2):229–232
16. Morris DM et al (2001) The reliability of the wolf motor function test for assessing upper extremity function after stroke. *Arch Phys Med Rehabil* 82(6):750–755
17. Bauder H, Taub E, Miltner WHR (2001) Behandlung motorischer Störungen nach Schlaganfall. Die Taubsche Bewegungsinduktionstherapie. Hogrefe, Göttingen
18. Wolf SL et al (2001) Assessing Wolf motor function test as outcome measure for research in patients after stroke. *Stroke* 32(7):1635–1639. doi:10.1161/01.str.32.7.1635
19. Schwickert L et al (2011) Robotic-assisted rehabilitation of proximal humerus fractures in virtual environments: a pilot study. *Z Gerontol Geriatr* 44(6):387–392
20. Fried LP et al (2001) Frailty in older adults: evidence for a phenotype. *J Gerontol A Biol Sci Med Sci* 56(3):146–156
21. Haidar SG et al (2004) Average versus maximum grip strength: Which is more consistent? *J Hand Surg Eur Vol* 29(1):82–84. doi:10.1016/j.jhsb.2003.09.012
22. Placzek JD et al (2004) Shoulder outcome measures: a comparison of 6 functional tests. *Am J Sports Med* 32(5):1270–1277
23. Bohannon RW (2006) Reference values for adult grip strength measured with a Jamar dynamometer: a descriptive meta-analysis. *Physiotherapy* 92(1):11–15
24. Whitall J (2006) Psychometric properties of a modified Wolf Motor Function test for people with mild and moderate upper-extremity hemiparesis. *Arch Phys Med Rehabil* 87(5):656–660
25. Cohen J (1968) Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol Bull* 70(4):213–220
26. Viera AJ, Garrett JM (2005) Understanding interobserver agreement: the kappa statistic. *Fam Med* 37(5):360–363
27. Pereira ND et al (2011) Reliability of the Brazilian version of the Wolf Motor Function Test in adults with hemiparesis. *Rev Bras Fisioter* 15(3):257–265
28. Nijland R et al (2010) A comparison of two validated tests for upper limb function after stroke: the Wolf Motor Function Test and the Action Research Arm Test. *J Rehabil Med* 42(7):694–696
29. George D (2005) PSS for Windows step-by-step: a simple guide and reference, 13.0 update, 6th edn. Pearson Education, Upper Saddle River, p 400
30. Bogard K et al (2009) Can the Wolf Motor Function Test be streamlined? *Neurorehabil Neural Repair* 23(5):422–428
31. Wu C et al (2011) Assessing the streamlined Wolf Motor Function Test as an outcome measure for stroke rehabilitation. *Neurorehabil Neural Repair* 25(2):194–199
32. Taub E, Morris DM, Crago J (2011) Wolf Motor Function Test (WMFT) Manual. Birmingham: University of Alabama, CI Therapy Research Group
33. Gallo JJ (2006) Handbook of geriatric assessment, 4th edn. Jones & Bartlett, Sudbury, p 475

Manuskript 2:

Nerz C, Schwickert L, Becker C, Studier-Fischer S, Müßig JA, Augat P. Effectiveness of robotic-assisted training added to conventional rehabilitation in patients with humeral fracture early after surgical treatment: study protocol of a randomised controlled trial. *Trials* 2017. DOI 10.1186/s13063-017-2274-z

This is an open access article. Permission to reuse this article is not required.

STUDY PROTOCOL

Open Access



Effectiveness of robot-assisted training added to conventional rehabilitation in patients with humeral fracture early after surgical treatment: protocol of a randomised, controlled, multicentre trial

Corinna Nerz^{1*}, Lars Schwickert¹, Clemens Becker¹, Stefan Studier-Fischer², Janina Anna Müßig³ and Peter Augat^{3,4}

Abstract

Background: The incidence of proximal humeral fractures increases with age. The functional recovery of the upper arm after such fractures is slow, and results are often disappointing. Treatment is associated with long immobilisation periods. Evidence-based exercise guidelines are missing. Loss of muscle mass as well as reduced range of motion and motor performance are common consequences. These losses could be partly counteracted by training interventions using robot-assisted arm support of the affected arm derived from neurorehabilitation. Thus, shorter immobilisation could be reached. Thus far, this approach has been tested in only a few small studies. The aim of the present study is to examine whether assistive robotic training augmenting conventional occupational and physical therapy can improve functional shoulder outcomes.

Methods/design: Patients aged between 35 and 66 years with proximal humeral fracture and surgical treatment will be recruited at three different clinics in Germany and randomised into an intervention group and a control group. Participants will be assessed before randomisation and followed after completing an intervention period of 3 weeks and additionally after 3, 6 and 12 months. The baseline assessment will include cognition (Short Orientation-Memory-Concentration Test); level of pain in the affected arm; ability to work; gait speed (10-m walk); disability of the arm, shoulder and hand (Disabilities of the Arm, Shoulder and Hand Outcome Measure [DASH]); range of motion of the affected arm (goniometer measurement); visual acuity; and motor function of orthopaedic patients (Wolf Motor Function Test–Orthopaedic version [WMFT-O]). Clinical follow-up directly after the intervention will include assessment of disability of the arm, shoulder and hand (DASH) as well as range of motion and motor function (WMFT-O). The primary outcome parameter will be the DASH, and the secondary outcome parameter will be the WMFT-O. The long-term results will be assessed prospectively by postal follow-up. All patients will receive conventional occupational and physical therapy. The intervention group will receive additional robot-assisted training using the Armeo®Spring robot for 3 weeks.

Discussion: This study protocol describes a phase II, randomised, controlled, single-blind, multicentre intervention study. The results will guide and possibly improve methods of rehabilitation after proximal humeral fracture.

Trial registration: Clinicaltrials.gov, NCT03100201. Registered on 28 March 2017.

Keywords: Robot-assisted rehabilitation, Humeral fracture, Randomised controlled study

* Correspondence: corinna.nerz@rbk.de

¹Department of Clinical Gerontology, Robert-Bosch-Hospital, Auerbachstrasse 110, 70376 Stuttgart, Germany

Full list of author information is available at the end of the article



Background

The number of upper arm fractures is increasing in the aging workforce (52–66 years). The risk for proximal humeral fractures (PHFs) increases with age, starting at around the age of 45 years [1]. In 2008 approximately 370,000 emergency department (ED) visits occurred as a result of humeral fractures in the United States. Of these around 50% were due to PHF (184,300 ED visits), with double the number in females compared with males. In the age group 35–64 years, approximately 57,500 ED visits were recorded. About 275,000 ED visits due to PHF are predicted in 2030 in the United States. A similar trend in numbers is expected in Europe [1].

This relevant number in fractures of the upper extremity already places a high demand on surgical treatment and subsequent rehabilitation [2]. Currently, the existing physiotherapy and occupational therapy concepts are based on expert opinion. No robust evidence-based consensus exists for rehabilitation approaches after PHF [3, 4], which is somewhat surprising. Worldwide no randomised controlled trials (RCTs) have been published comparing rehabilitation protocols for PHFs. Thus, the optimal content, frequency, duration and intensity of rehabilitation after shoulder injuries remain unclear. Not surprisingly, authors of a recent systematic literature review found considerable variation and heterogeneity in rehabilitative treatment after PHF [4]. Researchers in several observational studies have criticized the inadequacy of currently available rehabilitation protocols for PHFs. The functional recovery of the fractured upper extremity was found to be disappointing in a German study [5]. One consistent problem is the long immobilisation of the affected arm of up to 6 weeks [6]. This leads to loss of muscle strength and results in loss of flexibility of the adjacent joints. A shorter time to recovery would be a prerequisite for an improved outcome after PHF [4].

In a pilot safety and feasibility study by our group, robot-assisted training using the Armeo®Spring device (Hocoma, Volketswil, Switzerland), which was originally designed for stroke rehabilitation [7], was shown to be safe and highly acceptable even among older patients. The pre-post measurements indicated a potential improvement in muscle strength, flexibility and motor control [8]. Assistive arm support was used to reduce partial loading of the affected arm and increase the movement request at an early stage after surgery. Thus, robot-assisted training could be a valuable option in orthopaedic fracture rehabilitation. However, evidence for effectiveness so far is confined mainly to the field of neurology, particularly in stroke patients. Literature on fracture rehabilitation using robot-assisted training is extremely scarce. Besides

our pilot study, Padilla-Castaneda et al. [9] evaluated the use of robot-assisted training after forearm and elbow fracture. They showed that the duration and number of repetitions of exercises can be increased by active assistive training.

Although preliminary, the promising findings of these pilot studies remain to be confirmed in an RCT. Hence, the aim of this study is to test whether additive robot-assisted training in patients with PHF using the Armeo®Spring device is superior to conventional occupational and physical therapy alone.

Methods/design

Study design

A single-blind, multicentre intervention RCT will be conducted. Robot-assisted training using the Armeo®Spring device added to conventional occupational and physical therapy will be compared with a control group receiving only conventional occupational and physical therapy.

Randomisation

Participants will be randomised into one of two arms via block randomisation in each of the three centres, according to a predefined allocation sequence, after assessing inclusion and exclusion criteria. A randomisation list for each centre was created using the Internet-based Sealed Envelope software (<https://sealedenvelope.com/>; Sealed Envelope, London, UK). Assignment to the intervention and control groups will be performed by study nurses in each centre who will be blinded to interventions; unblinding will not be permissible. Anonymised data will be entered into a central database (Research Electronic Data Capture [REDCap]), with data entry in agreement with the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use standards for good clinical practice, see also SPIRIT procedure in supplementary material; Additional file 1: SPIRIT checklist and Additional file 2: Schedule of enrolment, interventions, and assessments) at the coordinating centre.

Participants will undergo clinical and quantitative assessments at one of the three sites: Robert-Bosch-Hospital Stuttgart, BG Trauma Centre Murnau or BG Trauma Centre Ludwigshafen. These three sites were chosen as comparators because they are the three largest clinical sites in Germany that treat PFH, do research or cooperate with an institute that conducts research, and possess Armeo®Spring devices. All procedures have been approved by the ethics committee of the University of Tübingen (381/2015BO1) and the regional medical association of Rheinland-Pfalz (837.519.15) and are in accord with the Declaration of Helsinki. Informed consent will be obtained from all participants (Additional file 3: Consent form). All participants will be insured. No data monitoring committee was consulted.

Sample size

Adults in the working population aged between 35 and 66 years will be recruited at all three sites. In a pilot study [8], our group was able to measure an improvement in function with the Disabilities of the Arm, Shoulder and Hand Outcome Measure (DASH) score (median 47 [mean 46 ± 20 SD] to median 34 [mean 32 ± 15 SD]). As a result of a power analysis based on these results, 26 patients per group will be required to achieve 80% power at a significance level of $\alpha = 0.05$. With an expected drop-out rate of 10%, a total of 60 patients should be recruited, comprising 20 patients at each of the 3 sites.

Inclusion criteria for both groups will be surgical treatment of the shoulder joint after upper arm fracture classified as AO Foundation/Association for the Study of Internal Fixation (Arbeitsgemeinschaft Osteosynthese; AO) type 11 [10]. Participants will be included between the fourth and seventh weeks after surgery. The participant population will be stratified by fracture location (subcapital/capital) and handedness.

Exclusion criteria will be determined during baseline assessment. Participants with limited cognition as defined by a score < 10 points on the Short Orientation-Memory-Concentration Test (SOMC), inadequate level of pain during movement of the affected shoulder joint (pain score > 5 on a pain visual analogue scale [VAS]), strongly limited vision or hearing, heart failure (New York Heart Association stage III-IV), chronic obstructive pulmonary disease with Global Initiative for Chronic Obstructive Lung Disease (GOLD) stage III-IV, walking speed < 0.8 m/second, isolated tuberculum majus fracture of the humerus (AO 11, A1), or fractures with involvement of the glenoid cavity, as well as double fractures or injury of the plexus or the axillary nerve, will be excluded. All participants will give their written informed consent to the study administrators at each centre prior to the clinical baseline assessment.

Clinical assessment

The estimated total duration of clinical baseline and follow-up assessments will be 160–180 minutes. The clinical assessment will be carried out for all participants before the first potential training session (baseline) and after the last potential training session (follow-up 1 at 3–5 weeks after baseline). For analysis of sustainability, further postal follow-up assessments (2, 3 and 4) will be carried out 3, 6 and 12 months after the baseline assessment (Fig. 1). In these follow-up assessments, only the DASH and the ability to work (what job to what extent) will be evaluated. All assessments will be performed by trained assessors.

Clinical data such as sex, age, body mass index, handedness, date and type of fracture, and date of

surgical treatment will be recorded. Furthermore, the participant's occupation before and after the accident as well as the number of days between the accident and the return to work will be recorded. If the participant will not be able to return to work after the accident, the reason for being incapacitated will be assessed.

The SOMC is an assessment of cognitive impairment and will be applied in this study. It can differentiate between mild, moderate and severe cognitive deficits. The SOMC was derived from the longer Blessed scale [11] to enable it to be performed by a non-physician. The SOMC score highly correlates ($r = 0.92$) with the full scale and is nearly as sensitive as the longer version. Any error score of 0–6 is within normal cognitive limits [12].

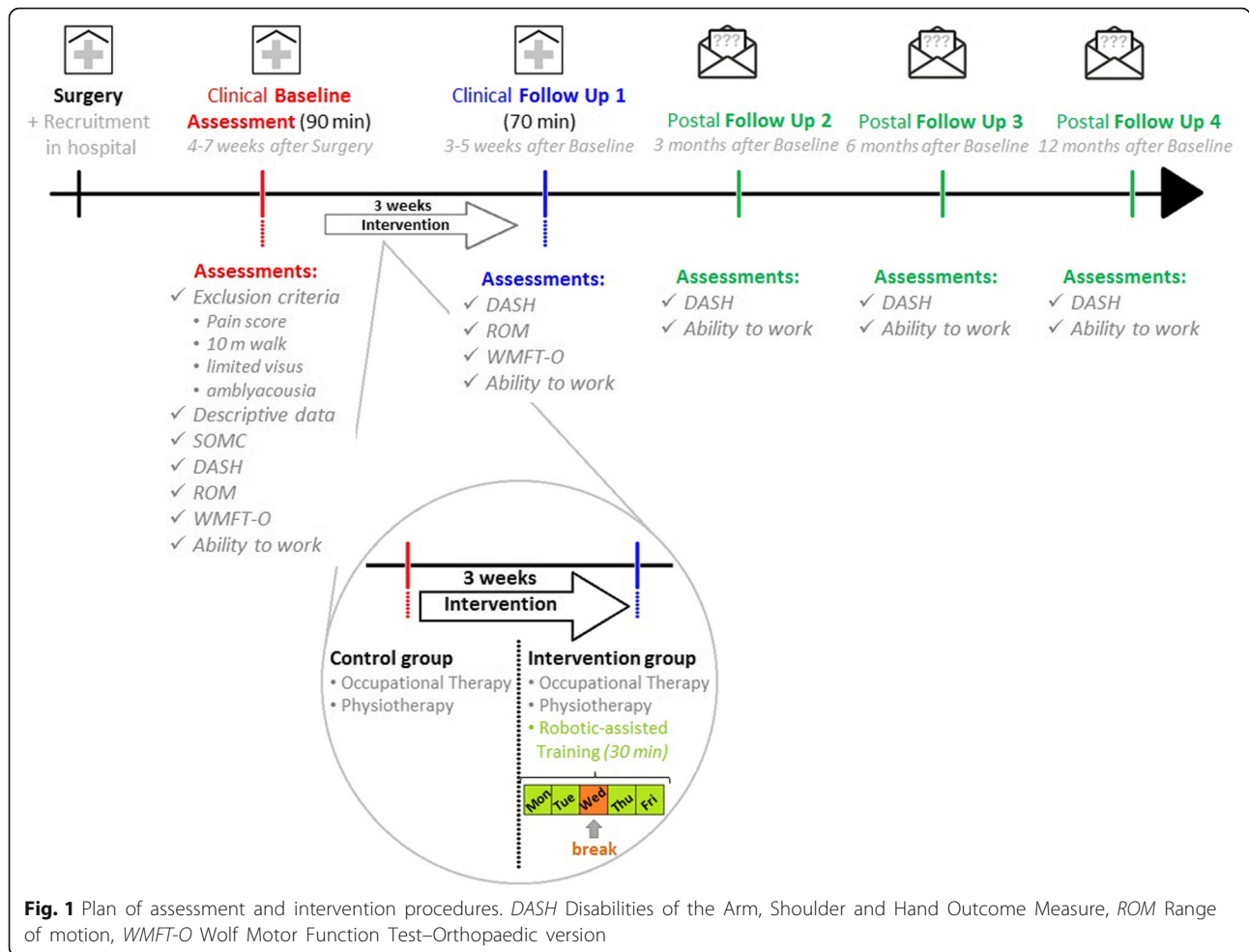
In this study the potential pain score in the shoulder during motion will be assessed by the patient and then documented by the therapist as a numerical value. Therefore, a VAS will be used, which was proven to be a valid method for determining pain in clinical routine [13]. As an assessment of overall motor function, the comfortable gait speed over a distance of 10 m will be assessed. The walking course will consist of a total of 14 m, 2 m for acceleration, 10 m for the speed measurement and 2 m for slowing down and stopping. The participants will be directed to walk at their own comfortable pace [14].

The flexibility (range of motion [ROM]) of the shoulder is measured using a goniometer according to the AO neutral-0 method in pain-free range. The maximum angle in elevation, retroversion, abduction and adduction as well as external rotation will be measured. The maximal grip strength will be assessed three times for both hands using a dynamometer (JAMAR Technologies/MSD, Londerzeel, Belgium) according to the American Society of Hand Therapists [15] guideline. The average of three measurements will be calculated for both hands.

In order to describe the extent to which the participants correctly follow medical advice and the therapy sessions, their adherence will be documented. This will be provided by recording the duration and frequency of additional training that contains conventional occupational and physical therapy (group or individualised) during the intervention.

Primary outcome

The primary outcome of self-perceived limitation of shoulder, arm and hand function will be assessed with the first module of the DASH [16], which is considered a valid and reliable test in clinical practice [17]. Using a total of 30 questions, we will assess the restriction relating to the function and activity of the



shoulder, arm and hand in daily living, as well as self-esteem and potentially existing symptoms of the shoulder, arm and hand, such as pain or prickle. The endpoint will be a score calculated from the ratings of the individual responses (1 is the best value, 5 the worst).

Secondary outcome

For objective proof of a functional improvement between the beginning and the end of the intervention, the Wolf Motor Function Test–Orthopaedic version (WMFT-O) [18] was conducted. The WMFT-O is an adapted version of the basic WMFT [19] with good inter- and intra-rater reliability. The WMFT-O includes 20 arm motion tasks of daily living that will be examined and evaluated in terms of time, functional capacity and quality of movement, arranged hierarchically and increasing from coarse movements

in the elbow and shoulder area to more complex and dexterous tasks in the fingers and the hand area.

Intervention

Both the intervention group and the control group received physical and occupational therapy for a period of 3 weeks; concomitant interventions were not restricted. The additional robot-assisted therapy commenced as early as 4 weeks after surgery and not later than 7 weeks after surgery. There are no evidence-based guidelines or requirements on the content, duration or intensity of rehabilitation after shoulder injuries. Typically the therapeutic content consists of exercises for stabilisation and mobilisation of the shoulder and the surrounding joints, as well as muscle strengthening (three to six times per week for 25–30 minutes). The obtained conventional physical and occupational therapy will be documented but not influenced. The frequency and duration of

individual training schedules, as well as adverse events, will be recorded in a protocol. Serious adverse events will be treated as criteria for discontinuing the interventions.

The intervention group received robot-assisted training in addition to the protocol intervention. The robot-assisted training will be performed using the clinically evaluated Armeo®Spring medical device. By means of sensory orthosis, arm movements will be supported and transmitted to a computer system and visually shown on a screen. Furthermore, audio feedback elements and augmented performance feedback will be used for guidance. The Armeo®Spring device encourages patients with motivating game-like exercises in an inspiring virtual environment. The sensor-based computer system calculates game kinetics for 3D animation of arm functions, including grasping and goal-orientated functional reaching, as well as object manipulation (e.g., throwing a Frisbee). The mechanism to support the weight of the arm will be increased or decreased gradually to adapt to the ability of the user. Thus, the participant will have the ability to perform targeted and controlled movements in accordance with medical safety standards, including the movement limit as suggested by the operating physician.

Preceding every therapy unit, saddle height and joint positioning of the functional segment parts of the exoskeleton will be individually adjusted to the upper extremity proportion of the affected arm of the participant. With the participant in an upright seating position with the spine on the back rest, the participant's arm will be attached to the exoskeleton with Velcro belts. A calibration of the ROM using the orthosis will be performed by the participants and assessed to calculate an individual game-specific motion envelope. This calibration will help to configure a safe environment for shoulder movements of the participants and will prevent overload and excessive demands. All participants of the intervention group at all three sites ($n = 30$) will receive a 3-week robotic intervention in a standardised training setting with supervised exercise four times per week for 30 minutes per session.

The normative levels of exercise capacity are inspired by a previous pilot study [8], and a rough guideline was created in consensus with the three centres. To create this guideline the 17 existing exercises were categorised according to the direction of the movement. Then 12 of the 17 existing exercises were selected for training of patients after PHF on the basis of the categorised movement directions. These 12 exercises are illustrated and described in Fig. 2 and Table 1.

A plan for the 3-week intervention was created in which the exercises gradually increased from 1D up to 3D movements. Each exercise will take 1–3 min, depending on the task and level of difficulty chosen. Each training unit will start with exercise in a pre-defined order, as shown in Fig. 3, whereas progression will depend on the participant's current physical performance ability, which will be decided by the therapist. Progression within one unit will be achieved by gradually increasing the level of difficulty and the number of repetitions of a single exercise and thus the duration of the exercise within one training unit. Progression of shoulder function over the intervention period will be achieved by gradually increasing dimensionality of the exercise and by performing different exercises, which require the participant to use more joint segments.

Statistics

The centre and the number of treatment units will be regarded as possible covariates for the primary outcome. All data assessed will be described descriptively (mean, SD, 95% CI and median for metric variables; frequency and percentile for categorical variables). After checking the data for normal distribution by comparing the distribution histograms and Shapiro-Wilk test, the group comparison will be carried out with the t test or the rank-based Wilcoxon test. For the primary outcome analysis of covariance with study centre, number of therapy units (to take account of the training inhomogeneity) and injury severity (to take account of the patient's homogeneity) as covariates will be performed. The secondary outcome (WMFT) will be analysed to generate study hypotheses using the methods described above. Inferential statistics will follow an intention-to-treat approach with mean imputation or regression imputation for all randomised patients with incomplete datasets if deemed necessary. Per-protocol analysis will be performed for evaluation of the efficacy of the intervention in patients with complete datasets and without protocol violations.

Discussion

PHFs are often associated with immobilisation of the affected arm, leading to functional loss, pain and psychosocial problems. The aim of rehabilitation is to reintegrate the patient as quickly as possible into daily routine and work after surgical treatment. This study protocol describes an experimental set-up allowing the analysis of an additive robot-assisted training in combination with conventional occupational and physical therapy in comparison with conventional occupational and physical therapy alone. The robot-assisted training will allow the patient to mobilise the affected arm earlier and thereby

Table 1 Description of the 12 exercises

Name of the task	Task description	Functional challenge
A. High flyer	Collect the coins while flying. Avoid obstacles like bombs and birds.	Flex and extend your shoulder to move up and down.
B. Rain mug	Catch the raindrops with your mug.	Abduct and adduct your shoulder horizontally to move the mug to the left and right sides.
C. Brick breaker	Destroy the bricks with the ball by controlling the moving platform. Making sure not to lose the ball.	Abduct and adduct your shoulder horizontally to move to the left and right sides.
D. Frisbee	Target the Frisbee at the balloons and bring the balloons to burst.	Flex, extend, adduct and abduct your shoulder. Bend and straighten your wrist to burst the balloons.
E. Helicopter rescue	Steer helicopters and rescue your robot friends off the roofs of buildings. Hover over them in order to pick them up.	Abduct and adduct your shoulder horizontally to move to the left and right sides. Rotate the forearm inward and outward to open and close the loading door of the helicopter.
F. Balloons	Catch the balloons while avoiding obstacles like bombs and birds.	Flex and extend your shoulder to move up and down. Abduct and adduct your shoulder horizontally to move to the left and right sides.
G. Fishing	Collect the fish and other aquatic animals, but no waste.	Flex and extend your shoulder to move up and down. Abduct and adduct your shoulder horizontally to move to the left and right sides.
H. Treasure Island	Take the key and roll the ball along the path over the bridge up to the treasure chest. Try to collect coins while doing this, and avoid the obstacles.	Flex and extend your shoulder to move up and down. Abduct and adduct your shoulder horizontally to move to the left and right sides.
I. Roll the ball	Move the ball and collect coins, avoiding the obstacles, like bugs and bombs.	Flex, extend, adduct and abduct your shoulder to move the ball.
J. Supermarket	Take the products, which are on the shopping list, from the shelves and put them in the trolley.	Flex, extend, adduct and abduct your shoulder to move the hand with which you can grab the products.
K. Diving	Collect all the coins in the underwater world while avoiding the bombs.	Flex, extend, adduct and abduct your shoulder to move the diver.
L. Farmer	Maintain your garden by using a variety of transaction types, from sowing to harvesting.	Flex, extend, adduct and abduct your shoulder to move the hand with which you can (e.g., watering the seed).

to lose less muscle mass and improve motor capacity. This will be measured objectively with the WMFT-O and subjectively with the self-completed DASH questionnaire. This will possibly allow for an earlier return to work.

To the best of our knowledge, this is the first controlled study in the field of rehabilitation of PHFs using robot-assisted training. The results of the study will give insight into treatment modalities after PHF and into restoration of shoulder function. This will be important for learning about ways to improve therapy in terms of duration, intensity, frequency and extent. New ways will be needed also to improve commitment to therapy. Therefore, the robot-assisted approach will be an option to offer more effective rehabilitation early after surgery.

Limitations

The sample size is relatively small and powered only for a functional endpoint (DASH). Social or economic endpoints such as return to work or cost-effectiveness of this intervention would require larger sample sizes and were not the purpose of this early-stage RCT. Furthermore, the set-up with three different associated centres will lead to some difficulties in the standardisation of the

assessments. It is expected that different observers and therapists in the three institutions have subjective differences in the accomplishment of the tests. Assessment of the WMFT will be videotaped to standardise this as much as possible.

Heterogeneous treatment modalities will also exist to some extent for rehabilitation after PHF in the conventional occupational and physical therapy in terms of duration, intensity and frequency. Theoretically, participation in the robotic group could lead to increased or decreased motivation to participate in other types of exercises. Adherence will therefore be documented.

For dissemination the full and optimised training protocol will be provided via an online research platform. Also two videos, one about the instructions of the WMFT-O and one about the rating, are provided on YouTube (<https://www.youtube.com/watch?v=8yDcWK9Xrtw>, <https://www.youtube.com/watch?v=MsfrGauYHfc>).

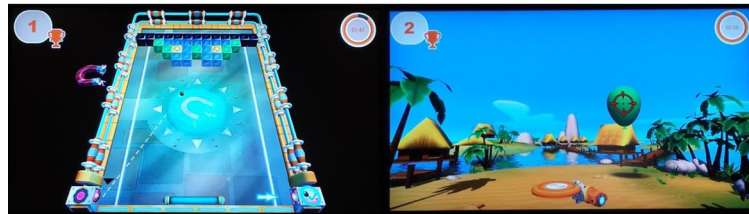
Conclusions

The results of the present RCT might help researchers in designing new therapeutic interventions in the rehabilitation of humeral fractures.



a: High Flyer

b: Rain Mug



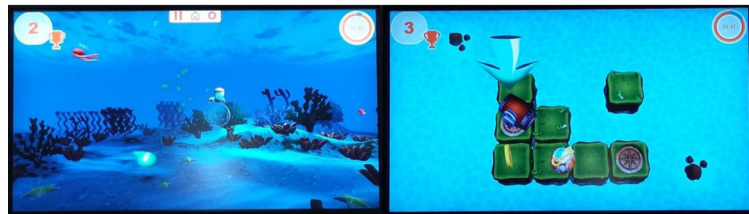
c: Brick Breaker

d: Frisbee



e: Helicopter Rescue

f: Balloons



g: Fishing

h: Treasure Island



i: Roll the Ball

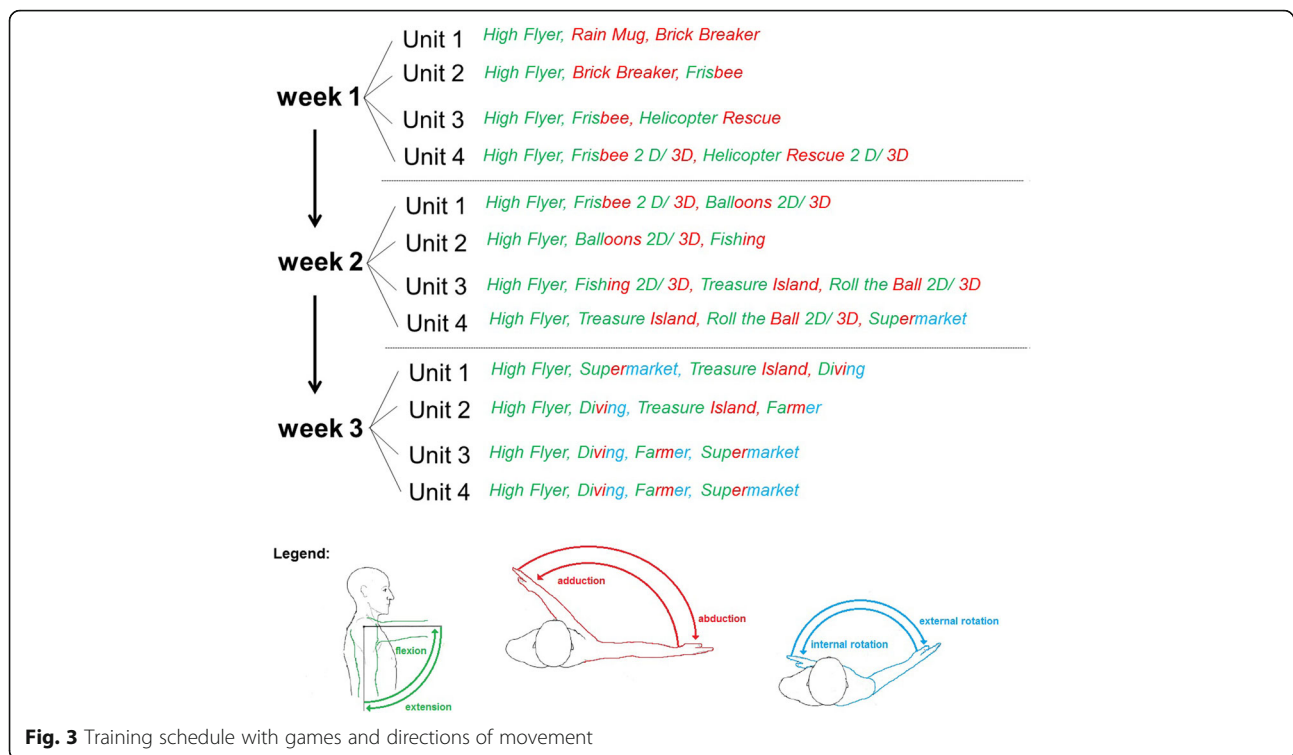
j: Supermarket



k: Diving

l: Farmer

Fig. 2 a-l Armeo®Spring games



Trial status

Patient recruitment was at 50% at the time of publication of this article.

Additional files

Additional file 1: Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) checklist. (DOC 121 kb)

Additional file 2: Schedule of enrolment, interventions, and assessments. (PDF 25 kb)

Additional file 3: Consent form. (DOCX 94 kb)

Abbreviations

AO: AO Foundation/Association for the Study of Internal Fixation (Arbeitsgemeinschaft Osteosynthese); DASH: Disabilities of the Arm, Shoulder and Hand Outcome Measure; ED: Emergency department; GOLD: Global Initiative for Chronic Obstructive Lung Disease; PHF: Proximal humeral fracture; RCT: Randomised controlled trial; REDCap: Research Electronic Data Capture; ROM: Range of motion; SOMC: Short Orientation-Memory-Concentration Test; VAS: Visual analogue scale; WMFT: Wolf Motor Function Test; WMFT-O: Wolf Motor Function Test–Orthopaedic version

Acknowledgements

The authors acknowledge the funding provided by the German Social Accident Insurance (project number 412.02-FR-0233).

Funding

The study is funded and supported by the Deutsche Gesetzliche Unfallversicherung (DGUV) - Alte Heerstrasse 111, 53757 St. Augustin. The DGUV had no influence on data assessment or the writing of the manuscript. Auditing procedures will be conducted annually.

Availability of data and materials

Data sharing is not applicable to this article, because no datasets were generated or analysed. All centres will have access to the anonymised trial data for analysis and publication of peer-reviewed journal articles.

Authors' contributions

CN, LS and PA were major contributors in the writing of the manuscript. All authors were involved in the methodological development of the study protocol and read and approved the final manuscript. The BG Trauma Centre Murnau will be responsible as coordinating centre, endpoint adjudication committee and data management team. All authors read and approved the final manuscript.

Ethics approval and consent to participate

For two of the three study sites (Robert-Bosch-Hospital Stuttgart and BG Trauma Centre Murnau), the study was approved by the ethics committee of the University of Tübingen (study number 381/2015B01). For the BG Trauma Centre Ludwigshafen, the study was approved by the regional medical association of Rheinland-Pfalz (study number 837.519.15 [10287]). Potential protocol modifications will be communicated with these sites.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Department of Clinical Gerontology, Robert-Bosch-Hospital, Auerbachstrasse 110, 70376 Stuttgart, Germany. ²Department of Trauma and Orthopaedic Surgery, BG Trauma Centre Ludwigshafen, Ludwig-Guttman-Strasse 13, 67071 Ludwigshafen, Germany. ³Institute of Biomechanics, BG Trauma Centre

Murnau, Professor-Küntscher-Strasse 8, 82418 Murnau am Staffelsee, Germany. ⁴Institute of Biomechanics, Paracelsus Medical University, Strubergasse 21, 5020 Salzburg, Austria.

Received: 4 April 2017 Accepted: 19 October 2017

Published online: 06 December 2017

References

- Kim SH, Szabo RM, Marder RA. Epidemiology of humerus fractures in the United States: nationwide emergency department sample, 2008. *Arthritis Care Res.* 2012;64(3):407–14.
- Häussler B, Gothe H, Göl D, Glaeske G, Pientka L, Felsenberg D. Epidemiology, treatment and costs of osteoporosis in Germany—the BoneEVA Study. *Osteoporos Int.* 2007;18(1):77–84.
- Zech A, Hübscher M, Vogt L, Banzer W, Hänsel F, Pfeifer K. Neuromuscular training for rehabilitation of sports injuries: a systematic review. *Med Sci Sports Exerc.* 2009;41(10):1831–41.
- Handoll HH, Brorson S. Interventions for treating proximal humeral fractures in adults. *Cochrane Database Syst Rev.* 2015;11, CD000434.
- Einsiedel T, Becker C, Stengel D, Schmelz A, Kramer M, Däxle M, Lechner F, Kinzl L, Gebhard F. Do injuries of the upper extremity in geriatric patients end up in helplessness? A prospective study for the outcome of distal radius and proximal humerus fractures in individuals over 65 [in German]. *Z Gerontol Geriatr.* 2006;39(6):451–61.
- Hodgson S. Proximal humerus fracture rehabilitation. *Clin Orthop.* 2006; 442:131–8.
- Liao WW, Wu C, Hsieh YW, Lin KC, Chang WY. Effects of robot-assisted upper limb rehabilitation on daily function and real-world arm activity in patients with chronic stroke: a randomized controlled trial. *Clin Rehabil.* 2012;26(2):111–20.
- Schwickert L, Klenk J, Stähler A, Becker C, Lindemann U. Robotic-assisted rehabilitation of proximal humerus fractures in virtual environments: a pilot study. *Z Für Gerontol Geriatr.* 2011;44(6):387–92.
- Padilla-Castaneda MA, Sotgiu E, Frisoli A, Bergamasco M, Orsini P, Martiradonna A, Olivieri S, Mazzinghi G, Laddaga C. A virtual reality system for robotic-assisted orthopedic rehabilitation of forearm and elbow fractures. In 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems, Tokyo, Japan, 3–7 Nov 2013. <http://ieeexplore.ieee.org/document/6696548/>. Accessed 21 Feb 2017.
- Baierlein SA. *Frakturklassifikationen*. Stuttgart: Georg Thieme Verlag; 2011.
- Blessed G, Tomlinson BE, Roth M. The association between quantitative measures of dementia and of senile change in the cerebral grey matter of elderly subjects. *Br J Psychiatry.* 1968;114(512):797–811.
- Katzman R, Brown T, Fuld P, Peck A, Schechter R, Schimmel H. Validation of a short Orientation-Memory-Concentration Test of cognitive impairment. *Am J Psychiatry.* 1983;140(6):734–9.
- Freyd M. The Graphic Rating Scale. *J Educ Psychol.* 1923;14(2):83–102.
- Flansbjerg UB, Holmbäck AM, Downham D, Patten C, Lexell J. Reliability of gait performance tests in men and women with hemiparesis after stroke. *J Rehabil Med.* 2005;37(2):75–82.
- American Society of Hand Therapists. *Clinical assessment recommendations*. 2nd ed. Chicago: American Society of Hand Therapists; 1992.
- Germann G, Harth A, Wind G, Demir E. Standardisation and validation of the German version 2.0 of the Disability of Arm, Shoulder, Hand (DASH) questionnaire [in German]. *Unfallchirurg.* 2003;106(1):13–9.
- Desai AS, Dramis A, Hearnden AJ. Critical appraisal of subjective outcome measures used in the assessment of shoulder disability. *Ann R Coll Surg Engl.* 2010;92(1):9–13.
- Oberle C, Becker C, Schölch S, Lenz JU, Studier-Fischer S, Augat P, Schwickert L. Inter- and intra-rater reliability of an adapted Wolf Motor Function Test for older patients with shoulder injuries. *Z Gerontol Geriatr.* 2017. doi:10.1007/s00391-017-1226-2.
- Wolf SL, Catlin PA, Ellis M, Archer AL, Morgan B, Piacentino A. Assessing Wolf Motor Function Test as outcome measure for research in patients after stroke. *Stroke.* 2001;32(7):1635–9.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit



Manuskript 3:

Nerz C, Schwickert L, Schölch S, Gordt K, Nolte PC, Kröger I, Augat P, Becker C. Inter-rater reliability, sensitivity to change and responsiveness of the orthopaedic Wolf-Motor-Function-Test as functional capacity measure before and after rehabilitation in patients with proximal humeral fractures. *BMC Musculoskeletal Disorders* 2019. DOI 10.1186/s12891-019-2691-0


This is an open access article. Permission to reuse this article is not required.

RESEARCH ARTICLE

Open Access



Inter-rater reliability, sensitivity to change and responsiveness of the orthopaedic Wolf-Motor-Function-Test as functional capacity measure before and after rehabilitation in patients with proximal humeral fractures

Corinna Nerz^{1*} , Lars Schwickert¹, Sabine Schölch¹, Katharina Gordt², Philip-Christian Nolte³, Inga Kröger⁴, Peter Augat^{4,5} and Clemens Becker¹

Abstract

Background: The incidence of proximal humeral fractures (PHF) increased by more than 30% over the last decade, which is accompanied by an increased number of operations. However, the evidence on operative vs. non-operative treatment and post-operative treatments is limited and mostly based on expert opinion. It is mandatory to objectively assess functional capacity to compare different treatments. Clinical tools should be valid, reliable and sensitive to change assessing functional capacity after PHFs. This study aimed to analyse inter-rater reliability of the videotaped Wolf-Motor-Function-Test-Orthopaedic (WMFT-O) and the association between the clinical WMFT-O and the Disability of the Arm, Shoulder and Hand (DASH) and to determine the sensitivity to change of the WMFT-O and the DASH to measure functional capacity before and after rehabilitation in PHF patients.

Methods: Fifty-six patients (61.7 ± 14.7 years) after surgical treatment of PHF were assessed using the WMFT-O at two different time points. To determine inter-rater reliability, the videotaped WMFT-O was evaluated through three blinded raters. Inter-rater agreement was determined by Fleiss' Kappa statistics. Pearson correlation coefficients were calculated to assess the association between the clinical WMFT-O and the video rating as well as the DASH. Sensitivity to change and responsiveness were analysed for the WMFT-O and the DASH in a subsample of forty patients (53.8 ± 1.4 years) who were assessed before and after a three week robotic-assisted training intervention.

Results: Inter-rater agreement was indicated by Fleiss' Kappa values ranging from 0.33–0.66 for functional capacity and from 0.27–0.54 for quality of movement. The correlation between the clinical WMFT-O and the video rating was higher than 0.77. The correlation between the clinical WMFT-O and the DASH was weak. Sensitivity to change was high for the WMFT-O and the DASH and responsiveness was given. In comparison to the DASH, the sensitivity to change of the WMFT-O was higher.

(Continued on next page)

* Correspondence: corinna.nerz@rbk.de

¹Department for Clinical Gerontology, Robert-Bosch-Hospital, Auerbachstr. 110, 70376 Stuttgart, Germany

Full list of author information is available at the end of the article



(Continued from previous page)

Conclusion: The overall results indicate that the WMFT-O is a reliable, sensitive and responsive instrument to measure more objectively functional change over time in rehabilitation after PHF. Furthermore, it has been shown that video assessment is eligible for studies to ensure a full blinding of raters.

Trial registration: Clinicaltrials.gov, [NCT03100201](https://clinicaltrials.gov/ct2/show/study/NCT03100201). Registered on 28 March 2017. The trial was retrospectively registered.

Keywords: Wolf motor function test, Reliability, Sensitivity to change, Responsiveness, Orthopaedic assessment, Shoulder function, Proximal humeral fracture

Background

The individual and societal burden of musculoskeletal injuries, in particular of bone fractures remains to be underestimated. Proximal humeral fractures (PHF) are among the leading causes of functional impairment in patients after trauma resulting in limitations in basic, instrumental and advanced activities of daily living. PHFs and wrist fractures are recognized as the most common fractures of the upper extremities accounting for more than 20% of hospital admissions caused by a fracture [1]. In patients over 40 years of age, the proportion of PHFs increases to 76% [2]. Since 2000 in Germany the incidence of PHFs has risen from 178 to 246/100.000 inhabitants/year [3]. In addition, an analysis from Bauer and colleagues showed that one third of the patients are still integrated into the work process [4]. Due to the demographic change, a further increase in the number of PHFs is expected [5, 6]. This will lead to a significant increase in PHFs requiring operative or non-operative treatment and post-trauma hospitalization and rehabilitation.

To date, there is no robust evidence-based consensus on rehabilitation after PHF regarding standardisation of content, duration, intensity or frequency [7–9]. One essential requirement to perform controlled studies on surgical and rehabilitation interventions is the availability of objective, reliable and valid assessments. If possible, these assessment tools should be blinded to treatment allocation. To assess functional capacity and task performance after PHE, at least two types of measures are required: patient-reported outcomes using questionnaires to assess activities of daily living and a supervised clinical-based assessment to measure functional capacity of the patients [10, 11]. The Disability of the Arm, Shoulder and Hand (DASH) questionnaire is the most commonly used questionnaire for assessing activities of daily living after shoulder and arm injuries [10]. A clinically administered assessment of the functional capacity including the quality of movement of PHF patients is the Wolf-Motor-Function-Test-Orthopaedic (WMFT-O) which has previously been assessed regarding re-test reliability, inter-rater reliability, and internal consistency [12]. One further property required of an outcome measurement is the sensitivity to change. It is understood as the ability to describe changes occurring during a treatment or observational

period. The particular meaning of this property is described by the fact that positive changes in a given period represent the classic therapeutic goal [13]. Beyond assessing functional change in clinical state over time with sufficient sensitivity to change [14–16] clinical-based assessment tools also need a high responsiveness to decide if a change over time is clinically meaningful [17, 18]. There are no studies that examined the sensitivity to change of the WMFT or the WMFT-O which shows the need to consider also the change over time of the functional capacity and the quality of movement of the WMFT-O.

The aims of this study were 1st to test the inter-rater reliability of the videotaped WMFT-O, 2nd to describe the correlation of the functional capacity assessed by the WMFT-O and the activities of daily living from a patient perspective assessed by the DASH questionnaire, and 3rd to describe the sensitivity to change and the responsiveness of the WMFT-O and the DASH in a group of patients with PHFs.

Methods

Patients

For testing inter-rater reliability of the videotaped WMFT-O two patient populations were assessed. The first sample was a group of sixteen patients with an age range from 75 to 90 years with surgical treatment after PHF [12]. Due to the funding guidelines of the sponsor of the study (“Deutsche Gesetzliche Unfallversicherung (DGUV)”) only persons up to 69 could be included into the intervention study. This decision is based on the study approach to be able to quickly re-integrate patients after PHF into everyday work [19]. The second group consisted of 40 patients with an age range from 34 to 69 years with surgical treatment after PHF participating in a randomised controlled trial to measure the effectiveness of robot-assisted training added to conventional rehabilitation [19]. The proximal humeral fractures of both patient groups were surgically fixed by plate osteosynthesis, screw fixation, endoprostheses or humeral nails. The second patient group was recruited at three different clinical sites in Germany and patients were randomised into an intervention group and a control group. They were assessed before randomisation (baseline) and after completing an intervention period of

3 weeks (reassessment). At baseline, cognition was assessed by the Short Orientation-Memory-Concentration Test [20] as well as visual acuity, gait speed (10-m walk [21]), level of pain in the affected arm, ability to work, disability of the arm, shoulder and hand (DASH [22]), range of motion of the affected arm (goniometer measurement [23]), and motor function of the affected arm and shoulder (WMFT-O; [12]). Clinical reassessment directly after the intervention assessed disability of the arm, shoulder and hand (DASH) as well as range of motion and functional capacity (WMFT-O). The WMFT-O was videotaped at both time points. To analyse the sensitivity to change and the responsiveness of the WMFT-O the second subsample with 40 patients was analysed. Both studies were approved by the ethical committee of the University of Tübingen and are in agreement with the Declaration of Helsinki. All participants gave written informed consent.

Outcome measures

The **WMFT-O** is an adapted version of the basic WMFT [24] with good clinical inter- and intra-rater reliability. The modified version for shoulder injuries was developed by our group [12]. The WMFT was developed to measure functional improvement between the beginning and the end of therapy. The WMFT-O is supervised and includes 20 arm motion tasks used in daily living that are evaluated in terms of functional capacity and quality of movement. The single items are progressing from coarse movements in the elbow and shoulder area to more complex and dexterous tasks in the fingers and the hand area. The endpoint is a total score calculated from the ratings of the functional capacity and the quality of movement (5 is the best value, 0 the worst). The total WMFT-O score ranges from 0 to 100 points with lower scores indicating greater disability (Additional file 1: Video tutorial Rating Scale of the WMFT-O: <https://youtu.be/WQUSl2XQJMY> and Additional file 2: Video tutorial performance instructions of the WMFT-O: https://youtu.be/K6g3Z_ibNa8). First a clinical rating was performed by pre-trained assessors. As a next step each videotaped WMFT-O (baseline and reassessment, $n = 112$) was assessed by three out of five pre-trained raters (4 physiotherapists, 1 occupational therapist).

The first module of the **DASH questionnaire** was used to measure self-perceived activity and limitation of shoulder, arm and hand function [22]. The DASH is considered as a valid and reliable non-supervised test [25–28]. The questionnaire consist of 30 questions to assess the restrictions related to the function and activity of the shoulder, arm and hand in daily living, as well as self-esteem and potentially existing symptoms of the shoulder, arm and hand, such as pain or prickle. The

endpoint is a total score calculated from the ratings of the individual responses (1 is the best value, 5 the worst). The total DASH score ranges from 0 to 100 points with higher scores indicating greater disability.

Data analysis

All statistics and outcome analyses were performed using RStudio software (Version 1.1.383). The inter-rater agreement of the video rating was determined by Fleiss' Kappa statistics with corresponding confidence intervals (CI). Pearson correlation coefficients were calculated to assess the association between the WMFT-O clinical rating and the video rating as well as between the WMFT-O clinical rating and the DASH questionnaire.

Sensitivity to change was defined as the ability of an instrument to respond to changes in the measured construct, regardless of whether the change is relevant or meaningful to the decider [14, 18]. In order to analyse the sensitivity to change of the WMFT-O, the change scores and the standardized effect size as well as the standardized response mean [29] were calculated. Change scores imply the delta of the results between the baseline- and reassessment. This score represents the extent of which a patient changes in performance in the corresponding test. The Wilcoxon test was calculated to compare the baseline and the reassessment score ($p > 0.01$). The standardized effect size was calculated by dividing the change score by the standard deviation of the baseline score [30]. The standardized response mean was calculated by dividing the change score by the standard deviation of that change score [31]. Values < 0.01 for the standardized effect size are considered as very small, < 0.2 as small, < 0.5 as medium, < 0.8 as large, < 1.2 as very large and < 2.0 as huge [32]. Husted and colleagues [29] interpreted the values of the standardized effect size and the standardized response mean considering the same benchmarks (< 0.2 for trivial, 0.2 to < 0.5 for small, 0.5 to < 0.8 for moderate, and 0.8 or greater for large). Cohen's threshold values are > 0.8 equates to large, > 0.5 to medium and > 0.2 to small effect sizes and intended for intervention studies, but are sometimes used to apply sensitivity to change of questionnaires [31].

Responsiveness was defined as the ability of an instrument to measure a meaningful or important change in a clinical state [14, 18] and is usually reported through the minimal important difference [33, 34]. For being considered as important, a change score of a measure should equal or exceed its minimal important difference estimate. Minimal important difference values were calculated using two commonly used effect size estimates in the literature: 0.3^* standard deviation of the baseline score and 0.5^* standard deviation of the baseline score [34, 35].

Results

The study sample consisted of two groups. The first subsample included 16 older patients with a mean age of 81.4 years (range 75–90 years). The second subsample includes 40 younger patients with a mean age of 53.8 years (range 39–69 years). The subject characteristics are listed in Table 1.

Inter-rater reliability of the videotaped WMFT-O

Table 2 lists inter-rater reliability across all items of the videotaped WMFT-O. The Fleiss' Kappa values ranged between 0.35 and 0.67 (CI = 0.27 to 0.73) for the functional capacity and between 0.27 and 0.54 (CI = 0.21 to 0.60) for the quality of movement. Grip strength (task 15) was not considered since the task could not be analysed from the video recordings.

Correlation between the WMFT-O clinical and video rating and the DASH

Pearson correlation coefficients (r) of the WMFT-O clinical rating and the video rating were high for all three raters and highly significant (functional capacity and quality of movement $r \geq 0.82$, $p < 0.001$) (Fig. 1). After reducing missing or occluded video recordings, the Pearson correlation coefficient for the functional capacity could be calculated for 49 subjects and for 48 subjects for the quality of movement. Due to the uniform results of the three different raters, in Fig. 1 only the ratings of the first rater are shown.

The correlation between the WMFT-O clinical rating and the DASH questionnaire was significantly weak at baseline (functional capacity $r = -0.27$ and quality of movement $r = -0.32$, $p < 0.05$) (Fig. 2). The Pearson correlation coefficient for the DASH could be calculated for 56 subjects.

Sensitivity to change and responsiveness

The sensitivity to change and the responsiveness indices of the WMFT-O and DASH are listed in Table 3. For the WMFT-O the clinical rating was carried out locally in the individual study centres by one trained rater and the DASH was completed by the subject himself on site. The baseline and the reassessment scores for the WMFT-O functional capacity and quality of movement as well as the DASH differ significantly from each other ($p < 0.01$). The WMFT-O and the DASH demonstrated large standardized effect sizes, ranging between 0.8 and 0.9 (Table 3). Large standardized response means were also obtained for the WMFT-O functional capacity, quality of movement and the DASH. The standardized response mean for the DASH was slightly less sensitive to change when compared to the WMFT-O functional capacity and quality of movement. The minimal important differences for the WMFT-O functional capacity ranged between 5.0 (0.3* standard deviation of the baseline score) and 8.3 (0.5* standard deviation of the baseline score) and for the WMFT-O quality of movement between 6.3 (0.3* standard deviation of the baseline score) and 10.5 (0.5* standard deviation of the baseline score) and are comparable to the minimal important differences of the DASH (Table 3).

Discussion

Our study on patients with fractures of the proximal humerus demonstrated high sensitivity to change and good responsiveness of the orthopaedic Wolf-Motor-Function-Test indicating its usefulness as a functional capacity assessment tool in operative and rehabilitation studies.

This study also found moderate inter-rater reliability for the videotaped version of the WMFT-O (Table 2). According to Landis and Koch [36] the calculated Fleiss'

Table 1 Participant demographics

Characteristic	1st Subsample	2nd Subsample	Total sample
Subjects (n)	16	40	56
Gender			
Female (%)	13 (81.2%)	25 (62.5%)	38 (67.9%)
Age (years)			
Mean (SD)	81.4 (1.1)	53.8 (1.4)	61.7 (14.7)
Minimum	75	39	39
Maximum	90	69	90
Affected Arm			
Right (%)	10 (62.5%)	17 (42.5%)	27 (48.2%)
Arm Dominance ^a			
Left (%)	1 (6.3%)	5 (12.5%)	6 (10.9%)
Right (%)	15 (93.7)	34 (85.0%)	49 (89.1%)
Dominant arm affectedd (%)	9 (56.3%)	17 (42.5%)	26 (47.3%)

n Number, SD Standard deviation; ^aDominant arm could not be determined in one subject.

Table 2 Inter-rater reliability across all items of the video WMFT-O; $n = 56$ ($p < 0.005$)

Task	Functional Capacity	Quality of Movement
	Inter-rater reliability Fleiss' Kappa (95% CI)	Inter-rater reliability Fleiss' Kappa (95% CI)
1. Forearm to table lateral	0.41 (0.35–0.48)	0.30 (0.24–0.37)
2a/b. Forearm to box 15/30 cm lateral	0.47 (0.41–0.53)	0.39 (0.34–0.44)
3. Forearm to box lateral with weight	0.64 (0.57–0.71)	0.54 (0.48–0.60)
4. Extend elbow lateral	0.49 (0.44–0.55)	0.46 (0.41–0.51)
5. Extend elbow lateral with weight	0.56 (0.50–0.62)	0.43 (0.38–0.49)
6. Hand to table frontal	0.60 (0.53–0.67)	0.38 (0.32–0.45)
7a/b. Hand to box 15/30 cm frontal	0.55 (0.49–0.61)	0.51 (0.46–0.57)
8. Hand to box frontal with weight	0.67 (0.60–0.73)	0.53 (0.48–0.59)
9. Reach and retrieve frontal	0.49 (0.43–0.55)	0.44 (0.37–0.50)
10. Lift can frontal	0.59 (0.52–0.65)	0.50 (0.44–0.56)
11. Lift pencil frontal	0.60 (0.53–0.67)	0.45 (0.38–0.51)
12. Lift paper clip frontal	0.51 (0.45–0.58)	0.39 (0.32–0.45)
13. Stack checkers frontal	0.50 (0.44–0.57)	0.34 (0.27–0.40)
14. Flip cards frontal ^a	0.49 (0.43–0.56)	0.37 (0.30–0.43)
16. Turn key in lock frontal	0.35 (0.27–0.42)	0.27 (0.21–0.33)
17. Fold towel frontal	0.49 (0.42–0.55)	0.42 (0.36–0.47)
18. Lift basket frontal	0.47 (0.41–0.53)	0.43 (0.37–0.48)

CI Confidence Interval, ^a task 15 (grip strength) was omitted.

Kappa values for the functional capacity and quality of movement can be interpreted as a fair to substantial agreement. The observed agreements were weaker than in a previous clinical study using a non-videtaped assessment of WMFT-O in similar patients [12]. Another publication investigating the reliability of the neurological based WMFT-N version [37] also had higher interrater agreement. The lower agreement of the videotaped measurement could have several reasons. Possibly the training session of the video raters were not enough. A further explanation for the lower agreement could also be an inadequate positioning of the chosen video camera and a different camera position might lead to better results. The performance of the tasks on such

videos (e.g. in task 16- “Turn key in lock frontal”, Fleiss’ Kappa for functional capacity =0.33 and quality of movement =0.27; Table 2) was not easy to identify which resulted in more diverse ratings between the raters. Moreover, an inadequate description of the correct end-position in task 1 (“Forearm to table lateral”, Fleiss’ Kappa functional for capacity =0.41 and quality of movement =0.30; Table 2) lead to different opinions about fulfilling or not fulfilling the task. A possible solution could be reducing the prescribed tasks to improve rater agreement. The tasks that were poorly rated could potentially be omitted and a future short version of the WMFT-O could be provided. According to Landis and Koch [36], an inter-rater reliability of 0.01 to 0.20 is considered as a

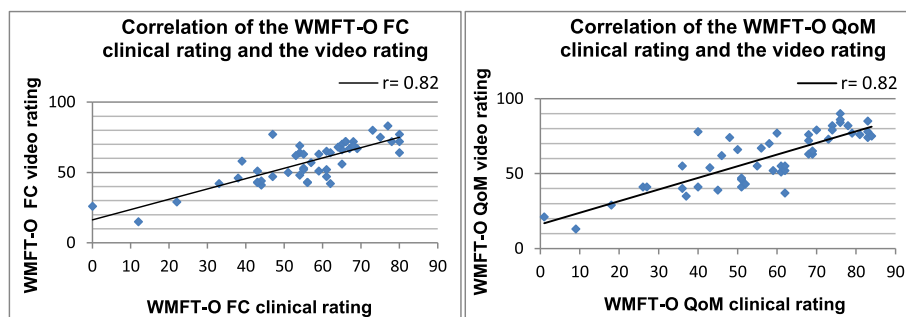


Fig. 1 Correlation of the WMFT-O functional capacity and quality of movement clinical rating and video rating (FC Functional capacity, QoM Quality of movement; Pearson correlation coefficient (r); $p < 0.001$)

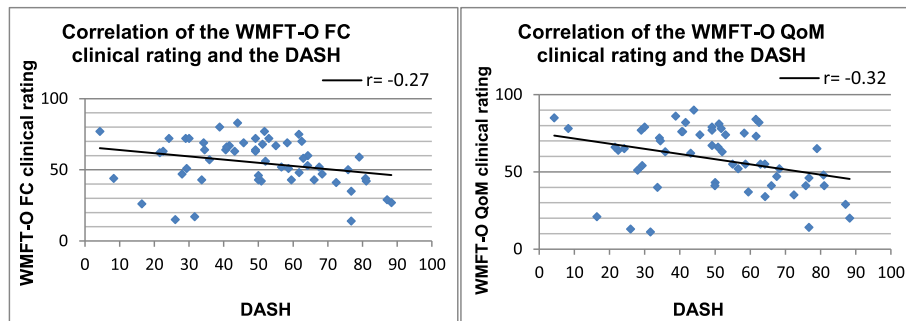


Fig. 2 Correlation of the WMFT-O functional capacity and quality of movement clinical rating and DASH (FC Functional capacity, QoM Quality of movement; Pearson correlation coefficient (r); $p < 0.05$)

slight agreement and 0.21 to 0.40 as a fair agreement. If these values were used as a basis to decide which items of the WMFT-O could be deleted in a short version, this would delete task 1 and 16 due to the low inter-rater reliability of the functional capacity and quality of movement as well as task 2ab, 12, 13 and 14 due to the low inter-rater reliability of the quality of movement (Table 2).

A third aspect is the calculation of different statistical measures. Inter-rater agreement for two different raters, as calculated in Oberle and colleagues 2018 [12] should be determined by weighted Cohen’s Kappa (K_w) statistics [38]. For three or more raters, Fleiss’ Kappas is recommended as the method of choice. The Fleiss’ Kappa assumes that the examiners were randomly selected from a group of available examiners. Cohen’s Kappa, on the other hand, assumes that examiners have been specifically selected and trained. Therefore, the probability of agreement in the Fleiss’ Kappa and the Cohen’s Kappa is estimated in different ways. In some cases, Fleiss’ Kappa in general may produce lower values even if the agreement is actually high as described before [39]. That in turn may be a possible explanation why the inter-rater reliability values of the video-based WMFT-O calculated by Fleiss’ Kappa (Table 2) were lower than the inter-rater reliability values of the clinically WMFT-O as reported by Oberle and colleagues [12] and calculated by Cohen’s Kappa.

We found a very strong correlation between the WMFT-O clinical baseline rating and the video baseline ratings for the functional capacity and a strong correlation

for the quality of movement according to the standards of Evans [40].

To assess treatment effects of interventions for musculoskeletal conditions, functional capacity and personal activity need to be evaluated. Therefore, the WMFT-O has to be augmented by other methods to evaluate levels of disability, activity and participation. A widely applied method is the DASH. It is expected that the correlation between activity levels and functional capacity is often less than anticipated. The correlation between the clinical-based measures (WMFT-O clinical rating of the functional capacity and the quality of movement) and the patient-reported questionnaires (DASH) at baseline was indeed weak [40]. One aspect is that patients do not return to their activity levels due to psychological problems such as insufficient self-efficacy. Other aspects are methodological problems. The DASH is not explicitly designed for the affected arm. This means that in instances where tasks are mostly carried out with the dominant hand (e.g. turning a key in a lock) the restriction in the non-dominant arm, shoulder and hand are not necessarily being captured through the DASH. This is only the case if the affected hand is also the dominant hand. This misjudgement could be avoided if care is taken that when answering the questions of the DASH, the assessment of the restriction always relates to the performance of the activity with the affected shoulder, arm or hand. If it is not possible to carry out the activity of daily living with the affected shoulder, arm or hand the patient must be able to imagine the execution of the

Table 3 Sensitivity to change and responsiveness of the clinical WMFT-O and DASH; $n = 40$

Parameter	Baseline Mean (SD)	Reassessment Mean (SD)	CS (SD)	SES	SRM	MID (0.35SD _b)	MID (0.5SD _b)
WMFT-O							
FC	54.3 (16.5)	69.2 (12.8)	14.9 (10.1)	0.9	1.5	5.0	8.3
QOM	57.7 (20.9)	76.1 (14.8)	18.4 (14.1)	0.9	1.3	6.3	10.5
DASH	51.9 (20.0)	36.7 (16.6)	-15.2 (12.6)	0.8	1.2	6.0	10.0

SD Standard deviation, CS Change score, SES Standardized effect size, SRM Standardized response mean, MID Minimally important difference, SD_b Standard deviation Baseline, WMFT-O Wolf-Motor-Function-Test-Orthopaedic, FC Functional capacity, QoM Quality of Movement.

activity of daily living and the possible restrictions as best as possible and then answer the question.

One study from Wu and colleagues [41] developed the streamlined WMFT which includes the performance rating of 6 timed tasks for neurological patients. They found a low effect size for the streamlined WMFT and the original WMFT [41]. In comparison to this study the WMFT-O had a large effect size for both the functional capacity and the quality of movement and can thus be regarded as being sensitive to change over time according to the published standards [29, 32]. This result was confirmed by the absolute values of the standardized response mean, which can be considered higher than the standardized effect sizes for both the functional capacity and the quality of movement. This is a relevant finding in terms of clinical use to provide a more objective and sensitive measure for assessing functional capacity and quality of movement of the upper extremities for patients in orthopaedic rehabilitation and may improve current assessments which currently are mostly subjective. Compared to the WMFT-O, the sensitivity to change of the DASH was lower. These findings indicate that the WMFT-O is a more sensitive outcome measure for assessing functional change over time through rehabilitation in patients with PHF. MacDermid and colleagues [42] and Westphal and colleagues [11] determined the sensitivity to change of the DASH for patients after wrist fractures. They found higher values for the effect size of the DASH in patients with wrist fractures after 0–3 months. After 3–12 months they observed to be somewhat lower [11]. This indicates that the standardized effect sizes and standardized response mean depend on the point of time of the baseline- and reassessment. In the first 12 weeks after baseline, a large treatment effect can be expected. The next three quarters might lead to further improvements but the effect sizes will be smaller. In our study, the baseline assessment was conducted approximately one month after surgery. In order to find out whether the WMFT-O can also be used to assess long-term therapy results, a future study should be carried out including longer therapy intervals assessing upper extremity functional capacity measured with the WMFT-O after three month. A similar relation was found for the DASH questionnaire. In conclusion, both the WMFT-O and the DASH are responsive and complimentary assessment instruments in order to measure the functional change in patients with PHF.

Conclusion

The WMFT-O is a responsive instrument to objectively measure patient functional change in younger and older patients with PHF. It showed a somewhat higher sensitivity to change in younger and older patients with PHF

compared to the DASH. For assessing treatment success of interventions and rehabilitation functional capacity could be measured by the WMFT-O but it should be augmented by another method (for example the DASH) to enable the evaluation of levels of disability, activity and participation. We recommend minor modifications of the camera positions and of the descriptions of some movement tasks for the videotaped WMFT-O. After a detailed training of the rater and taking into account these changes we also recommend the application of a video based assessment as this would allow robust blinding of assessors which often is not the case due to contamination of therapists and assessment staff. Furthermore a short version of the WMFT-O would be desirable.

Additional files

Additional file 1: Video tutorial Rating Scale of the WMFT-O. Video tutorial with detailed instructions how to rate the single items of the WMFT-O (functional capacity and quality of movement) in English. (MP4 6957 kb)

Additional file 2: Video tutorial performance instructions of the WMFT-O. Video tutorial with detailed instructions how to perform the single items of the WMFT-O in English. (MP4 21016 kb)

Abbreviations

CI: Confidence Intervals; CS: Change Score; DASH: Disability of the Arm, Shoulder and Hand; FC: Functional Capacity; MID: Minimally Important Difference; PHF: Proximal Humeral Fractures; QoM: Quality of Movement; r: Pearson correlation coefficients; SD: Standard Deviation; SD_b: Standard Deviation baseline; SES: Standardized Effect Size; SRM: Standardized Response Mean; WMFT-O: Wolf-Motor-Function-Test-Orthopaedic; Kw: Cohen's Kappa

Acknowledgments

The authors gratefully acknowledge C. Endress (BSc), E. Ermer (BSc) and R. Leonhardt (BSc) for their support in rating the videotaped WMFT-O.

Authors' contributions

CN, LS, CB and PA were major contributors in the writing of the manuscript. CN, LS, CB, PN and PA were involved in the methodological development of the study protocol. CN, SS, KG and IK helped with the recruitment and data collection. All authors read and approved the final manuscript.

Funding

The study is funded and supported by the German Social Accident Insurance ("Deutsche Gesetzliche Unfallversicherung (DGUV) - Alte Heerstrasse 111, 53757 St. Augustin", project number 412.02-FR-0233). The DGUV had no influence on data assessment or the writing of the manuscript.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

Both studies were approved by the ethical committee of the University of Tübingen (study number 381/2015BO1) and are in agreement with the Declaration of Helsinki. All participants gave written informed consent.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department for Clinical Gerontology, Robert-Bosch-Hospital, Auerbachstr. 110, 70376 Stuttgart, Germany. ²Network Aging Research (NAR), Heidelberg University, Heidelberg, Germany. ³Department of Trauma and Orthopaedic Surgery, BG Trauma Centre Ludwigshafen, Ludwigshafen, Germany. ⁴Institute for Biomechanics, BG Unfallklinik Murnau, Murnau, Germany. ⁵Institute for Biomechanics, Paracelsus Medical University, Salzburg, Austria.

Received: 12 March 2019 Accepted: 25 June 2019

Published online: 06 July 2019

References

- Benzinger P, Riem S, Bauer J, Jaensch A, Becker C, Büchele G, et al. Risk of institutionalization following fragility fractures in older people. *Osteoporos Int*. 2019;1–8. Epub ahead of print.
- Kara H, Bayir A, Ak A, Acar D, Akinci M, Degirmenci S. Trivial trauma induced bilateral proximal end Humerus fracture: two case reports. *J Case Rep*. 2013; 3(2):366–9.
- Statistisches Bundesamt. Krankenhausstatistik - Diagnosedaten der Patienten und Patientinnen in Krankenhäusern. Wiesbaden: Statistisches Bundesamt Deutschland; c2017 [cited 2019 Feb 25]. Available from: <https://www-genesis.destatis.de/genesis/online?sequenz=statistikTabellen&selectionname=23131>. German.
- Bauer M. Klinikinterne analyse epidemiologischer Aspekte der proximalen Humerusfraktur [dissertation]. Tübingen: Universität Tübingen; 2014. [cited 2019 Feb 18]. Available from: <https://publikationen.uni-tuebingen.de/xmlui/handle/10900/56031>. German
- Palvanen M, Kannus P, Niemi S, Parkkari J. Update in the epidemiology of proximal humeral fractures. *Clin Orthop* Januar. 2006;442:87–92.
- Kim SH, Szabo RM, Marder RA. Epidemiology of humerus fractures in the United States: nationwide emergency department sample, 2008. *Arthritis Care Res*. 2012;64(3):407–14.
- Zech A, Hübscher M, Vogt L, Banzer W, Hänsel F, Pfeifer K. Neuromuscular training for rehabilitation of sports injuries: a systematic review. *Med Sci Sports Exerc*. 2009;41(10):1831–41.
- Franke S, Ambacher T. Die proximale Humerusfraktur. *Obere Extrem*. 2012; 7(3):137–43 German.
- Handoll HH, Ollivere BJ, Rollins KE. Interventions for treating proximal humeral fractures in adults. *Cochrane Database Syst Rev*. c2012;12. Available from: <https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD000434.pub3/abstract>. [cited 2019 Feb 18].
- Vincent JL, MacDermid JC, King GJ, Grewal R. Validity and sensitivity to change of patient-reported pain and disability measures for elbow pathologies. *J Orthop Sports Phys Ther*. 2013;43(4):263–74.
- Westphal T. Reliability and responsiveness of the German version of the disabilities of the arm, shoulder and hand questionnaire (DASH). *Unfallchirurg*. 2007;110(6):548–52 German.
- Oberle C, Becker C, Schölch S, Lenz J-U, Studier-Fischer S, Augat P, et al. Inter-rater and intra-rater reliability of an adapted Wolf motor function test for older patients with shoulder injuries. *Z Für Gerontol Geriatr*. 2018;51(3): 293–300.
- Beaton DE. Understanding the relevance of measured change through studies of responsiveness. *Spine*. 2000;25(24):3192–9.
- Liang MH, Lew RA, Stucki G, Fortin PR, Daltroy L. Measuring clinically important changes with patient-oriented questionnaires. *Med Care*. 2002; 40(4 Suppl):145–51.
- Schmitt JS, Di Fabio RP. Reliable change and minimum important difference (MID) proportions facilitated group responsiveness comparisons using individual threshold criteria. *J Clin Epidemiol*. 2004;57(10):1008–18.
- Terwee CB, Dekker FW, Wiersinga WM, Prummel MF, Bossuyt PMM. On assessing responsiveness of health-related quality of life instruments: guidelines for instrument evaluation. *Qual Life Res Int J Qual Life Asp Treat Care Rehab*. 2003;12(4):349–62.
- VanSwearingen JM, Brach JS. Making geriatric assessment work: selecting useful measures. *Phys Ther*. 2001;81(6):1233–52.
- Liang MH. Longitudinal construct validity: establishment of clinical meaning in patient evaluative instruments. *Med Care*. 2000;38(9 Suppl):1184–90.
- Nerz C, Schwickert L, Becker C, Studier-Fischer S, Müßig JA, Augat P. Effectiveness of robot-assisted training added to conventional rehabilitation in patients with humeral fracture early after surgical treatment: protocol of a randomised, controlled, multicentre trial. *Trials*. 2017;18(1):589.
- Katzman R, Brown T, Fuld P, Peck A, Schechter R, Schimmel H. Validation of a short orientation-memory-concentration test of cognitive impairment. *Am J Psychiatry*. 1983;140(6):734–9.
- Flansbjerg U-B, Drake AM, Downham D, Patten C, Lexell J. Reliability of gait performance tests in men and women with hemiparesis after stroke. *J Rehabil Med*01-010100. 2005;37(2):75–82.
- Germann G, Harth A, Wind G, Demir E. Standardisierung und Validierung der deutschen version 2.0 des "disability of arm, shoulder, hand" (DASH)-Fragebogens zur outcome-Messung an der oberen Extremität. *Unfallchirurg*. 2003;106(1):13–9 German.
- Ryf C, Weymann A. Range of motion - AO Neutral-0 method : measurement and documentation. Stuttgart: Thieme; 1999. p. 116.
- Wolf SL, Catlin PA, Ellis M, Archer AL, Morgan B, Piacentino A. Assessing Wolf Motor function test as outcome measure for research in patients after stroke. *Stroke*. 2001;32(7):1635–9.
- Desai AS, Dramis A, Hearnden AJ. Critical appraisal of subjective outcome measures used in the assessment of shoulder disability. *Ann R Coll Surg Engl*. 2010;92(1):9–13.
- Offenbaecher M, Ewert T, Sangha O, Stucki G. Validation of a German version of the disabilities of arm, shoulder, and hand questionnaire (DASH-G). *J Rheumatol*. 2002;29(2):401–2.
- Dalton E, Lannin NA, Laver K, Ross L, Ashford S, McCluskey A, et al. Validity, reliability and ease of use of the disabilities of arm, shoulder and hand questionnaire in adults following stroke. *Disabil Rehabil*. 2017;39(24):2504–11.
- Raven EEJ, Haverkamp D, Sierveit IN, van Montfoort DO, Pöhl RG, Blankevoort L, et al. Construct validity and reliability of the disability of Arm, shoulder and hand questionnaire for upper extremity complaints in rheumatoid arthritis. *J Rheumatol*. 2008;35(12):2334–8.
- Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: a critical review and recommendations. *J Clin Epidemiol*. 2000;53(5):459–68.
- Kazis L, Anderson J, Meenan R. Effect sizes for interpreting changes in health status. *Med Care*. c 1989 27(3) [cited 2019 Feb 19]. Available from: insights.ovid.com.
- Liang MH, Fossel AH, Larson MG. Comparisons of five health status instruments for orthopedic evaluation. *Med Care*. 1990;28(7):632.
- Sawilowsky S. New effect size rules of thumb. *Theor Behav Found Educ Fac Publ*. 2009;8(2):597–99. Available from: https://digitalcommons.wayne.edu/coe_tbf/4
- Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol*. 2008;61(2):102–9.
- Eton DT, Yost KJ. Combining distribution- and anchor-based approaches to determine minimally important differences: the FACIT experience. *Eval Health Prof*. 2005;28(2):172–91.
- Cella D, Eton DT, Fairclough DL, Bonomi P, Heyes AE, Silberman C, et al. What is a clinically meaningful change on the functional assessment of cancer therapy-lung (FACT-L) questionnaire?: results from eastern cooperative oncology group (ECOG) study 5592. *J Clin Epidemiol*. 2002; 55(3):285–95.
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159–74.
- Pereira ND, Michaelsen SM, Menezes IS, Ovando AC, Lima RCM, Teixeira-Salmela LF. Reliability of the Brazilian version of the Wolf Motor function test in adults with hemiparesis. *Braz J Phys Ther*. 2011;15(3):257–65.
- Cohen J. Weighted kappa. *Psychol Bull*. 1968;70(4):213–20.
- Falotico R, Quatto P. Fleiss' kappa statistic without paradoxes. *Qual Quant*. 2015;49(2):463–70.
- Evans JD. Straightforward statistics for the behavioral sciences. Pacific Grove: Brooks/Cole Publishing; 1996.
- Wu CY, Fu T, Lin KC, Feng CT, Hsieh KP, Yu HW, et al. Assessing the streamlined Wolf Motor function test as an outcome measure for stroke rehabilitation. *Neurorehabil Neural Repair*. 2011;25(2):194–9.
- MacDermid JC, Richards RS, Roth JH. Distal radius fracture: a prospective outcome study of 275 patients. *J Hand Ther*. 2001;14(2):154–69.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Manuskript 4:

Kröger I, **Nerz C**, Schwickert L, Schölch S, Müßig JA, Studier-Fischer S, Nolte PC, Becker C, Augat P. Robot-assisted training after proximal humeral fracture: A randomised controlled multicentre intervention trial. *Clinical Rehabilitation*. 2021;35(2):242-252. DOI10.1177/0269215520961654 (*Original submission version*)

* *Geteilte Erstautorenschaft*

Gratis reuse in dissertation or thesis, including where the dissertation or thesis will be posted in any electronic Institutional Repository or database.

Robot-assisted training after proximal humeral fracture – a randomised controlled multicentre intervention trial

Inga Kröger^{1,4†}, Corinna Nerz^{2†}, Lars Schwickert², Sabine Schölch², Janina Anna Müßig^{1,4}, Stefan Studier-Fischer³, Philip-Christian Nolte³, Clemens Becker² and Peter Augat^{1,4}

†Inga Kröger and Corinna Nerz shared first authorship

¹Institute for Biomechanics, Berufsgenossenschaftliche Unfallklinik Murnau, Murnau, Germany

²Department for Clinical Gerontology, Robert-Bosch-Hospital, Stuttgart, Germany

³Department of Trauma and Orthopaedic Surgery, Berufsgenossenschaftliche Unfallklinik Ludwigshafen, Ludwigshafen, Germany

⁴Institute for Biomechanics, Paracelsus Medical University, Salzburg, Austria

Corresponding Author:

Corinna Nerz

Department for clinical gerontology, Robert-Bosch-Hospital,

Auerbachstr. 110, 70376 Stuttgart, Germany

Corinna.nerz@rbk.de

Contact details:

Inga Kröger: inga-kroeger@posteo.de; Corinna Nerz: Corinna.Nerz@rbk.de; Lars Schwickert:

Lars.Schwickert@rbk.de; Sabine Schölch: Sabine.Schoelch@rbk.de; Janina Anna Müßig:

janinamuessig@aol.de; Stefan Studier-Fischer: stefan.studier-fischer@bgu-ludwigshafen.de; Philip-

Christian Nolte: PhilipChristian.Nolte@bgu-ludwigshafen.de; Clemens Becker:

Clemens.Becker@rbk.de; Peter Augat: Peter.Augat@bgu-murnau.de

Abstract

Objective: To examine whether robotic-assisted training as a supplement to conventional therapy is safe, acceptable and improves function and patient reported outcome after proximal humeral fractures (PHF).

Design: Multicentre, assessor-blinded, randomised controlled prospective trial.

Setting: Three different rehabilitation hospitals in Germany.

Subjects: In total 928 PHF patients between 35-70 years were screened. 48 participants were included in the study (intervention group n= 23; control group n=25).

Intervention: The control group received conventional occupational and physiotherapy over 3 weeks, and the intervention group received additional 12 robot-assisted training sessions at the ARMEO®-Spring.

Main measures: Disabilities of the Arm, Shoulder and Hand Questionnaire (DASH), the Wolf Motor Function Test-Orthopaedic, active range of motion and grip strength were determined before and after intervention period. The DASH was additionally obtained postal 6 and 13 months following surgery.

Results: The mean age of participants was 55 ± 10 years and was similar in both groups ($p > 0.05$). The change in DASH as the primary endpoint in the intervention group after intervention was -15 (CI= 8-22), at follow-up 6 month -7 (CI= -2-16) at follow up 12 month -9 (CI= 1-16); in control group -14 (CI= 11-18), at follow-up 6 month -13 (CI= 7-19) at follow up 13 month -6 (CI= -3-14). No difference in the change was found between groups ($p > 0.05$). None of the follow-up time points demonstrated an additional benefit of the robotic therapy.

Conclusion: The additional robot-assisted therapy was safe, acceptable but showed no improvement in functional shoulder outcome compared to conventional therapy only.

Keywords: Robot-assisted rehabilitation, Humeral fracture, Occupational and Physiotherapy, Robotics, Patient Reported Outcomes

Introduction

Delayed recovery and incomplete restoration of shoulder function constitute social and economic burdens on the patients as well as economic burdens on the health system and society, since individuals affected by humeral fractures have a mean age of 55 years and are often still in their active working years [1,2]. The functional recovery of the upper arm after a humeral fracture remains disappointing [3]. Diverse problems, such as muscular atrophy, loss in range of motion, loss of flexibility of the adjacent joint and consequently diminished working ability and psychosocial problems arise during the long immobilisation period of the affected arm [3, 4]. To this date there are few studies concerning the optimal type of rehabilitation following proximal humeral fracture. The existing physiotherapy and occupational therapy concepts are generally based on expert opinion. No evidence-based consensus exists for rehabilitation approaches after proximal humeral fracture considering content, time, duration and frequency [5, 6]. A 2017 review included three studies that investigated the effect of electrotherapy, hydrotherapy and trigger point treatment on functional recovery after humeral fracture of which only trigger point treatment showed some improvement in outcome [7].

Robotics-supported training is gaining attention in clinical rehabilitation, particularly in neurorehabilitation [8]. Gijbels and colleagues showed functional benefits of using robot-assisted training in patients with multiple sclerosis [9]. Similarly, several studies have shown improvements in motor control and strength of the paretic upper limb in stroke patients after robot-assisted training [8]. Robotic training as treatment after fracture is rarely reported. Padilla-Castañeda and colleagues developed a robotic system with the aim of extending range of motion, strengthening and pain reduction for orthopaedic patients after elbow and wrist injuries. The system has only been tested on a small number of volunteers, but not in symptomatic patients [10]. A pilot study by Schwicker and colleagues demonstrated positive trends in clinical outcomes using the ARMEO®-Spring (Hocoma AG, Zurich, Switzerland) device in the rehabilitation of geriatric patients after proximal humeral fracture [6]. In summary, rehabilitation modalities in proximal humeral fracture are still suboptimally described and controlled studies are required [11].

The aim of this randomised controlled trial was to examine whether robotic-assisted training as a supplement to conventional therapy is safe, acceptable and improves function and patient-reported outcomes. The hypothesis was that the additional intervention using assistive robotic training (ARMEO®-Spring) over a period of three weeks is superior to conventional therapy and improves the functional outcome in patients with proximal humeral fracture [12].

Methods

The study was a multicentre, assessor-blinded, randomised controlled prospective trial. The study protocol has been previously published [12]. The trial was registered in the International Register for Clinical Trials (Clinicaltrials.gov, NCT03100201) and the German Register for Clinical Trials (DRKS00009990). All procedures have been approved by the ethics committee of the state medical association Rheinland-Pfalz (837.519.15 (10287)) and the ethics committee at the medical faculty of the Eberhard-Karls-University Tübingen (381/2015BO1) and are in accordance with the Declaration of Helsinki. The study was funded by the German Social Accident Insurance (DGUV) project FP-0233: "Robotic assisted training after humeral fracture – a randomised intervention study" and was conducted from February 2016 to December 2019 in three rehabilitation clinics in Germany (Robert-Bosch-Hospital Stuttgart; BG Unfallklinik Murnau and BG Unfallklinik Ludwigshafen).

Participants between 35 and 70 years of age with a surgically treated proximal humeral fracture (AO classification; Association for the Study of Internal Fixation: AO11; [13]) were screened and included between the fourth and seventh weeks after surgery at all three rehabilitation hospitals in this study. The fractures were surgically fixed by plate osteosynthesis, screw fixation, endoprotheses or humeral nails. Exclusion criteria were defined as limited cognition as defined by a score <10 points on the Short Orientation-Memory-Concentration Test [14], inadequate level of pain during movement of the affected shoulder joint (pain score >5 on a pain visual analogue scale [15], strongly limited vision or hearing, heart failure (New York Heart Association stage III-IV), chronic obstructive pulmonary disease with Global Initiative for Chronic Obstructive Lung Disease stage III-IV, walking speed <0.8 meter/second, isolated fractures of the tuberculum major, glenoid fractures, or multiple fractures and injuries of the plexus or axillary nerve [12]. All participants gave signed informed consent prior to enrolment.

Participants were randomised to a control group or an intervention group. Randomisation was stratified via block randomisation in each of the three centres (Stuttgart, Murnau, and Ludwigshafen), according to a predefined allocation sequence on the basis of the fracture location (capital or subcapital proximal humeral fracture) and the dominance of the fracture side as a stratification criterion. Fracture classification was carried out by the treating physician or assessor. A randomisation list for each centre was created using the Internet-based Sealed Envelope software (<https://sealedenvelope.com>; Sealed Envelope, London, UK).

All participants received the clinical baseline assessment between the fourth and seventh week after surgery and a clinical follow-up assessment after the three-week intervention period (7-12 weeks after surgery). Clinical baseline and follow-up included the first module of the patient-reported Disabilities of the Arm, Shoulder and Hand Outcome Questionnaire as the primary outcome measure and the Wolf Motor Function Test Orthopaedic version [16-18]. Further clinical assessments during baseline and follow up included a goniometer-based measurement of the active range of motion of the shoulder joint [19] and grip strength using the Jamar dynamometer [20]. The Disabilities of the Arm, Shoulder and Hand Outcome Questionnaire was additionally evaluated via a postal questionnaire at 6 months (6 +/-1 month after surgery) and 13 months (13 +/-2 month after surgery) postoperatively. All participants were provided with a therapy diary documenting the type and extent of conventional therapy during the three-week intervention period. In order to assess the safety of the intervention, any serious or other adverse events were documented. Drop-outs were documented to check whether these were due to a lack of acceptance of the intervention [12]. Study data were collected and managed using a central database (Research Electronic Data Capture [REDCap]; [21]), with data entry in agreement with the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use standards for good clinical practice.

The control group received during the three weeks intervention standard of care including occupational and physical therapy. The intervention group received in addition to standard of care, twelve 30-minute robotic-assisted therapy sessions using the ARMEO®-Spring. The ARMEO®-Spring is a medically certified robotic training device in which the arm is placed in a sensory-based exoskeleton which transmits the 3D motion to a computer system to visually show on a screen and to manipulate virtual objects in a game (Figure 1). The patient actively moves the arm, whereby the degree of assistance in lifting the arm can be individually adjusted to the ability of the user. Movement limitations prescribed by the operating physician and therapists were considered personalised to allow a training according to medical safety standards. The exergames challenge the patient to perform various functional tasks such as filling a shopping basket or moving objects. By use of the ARMEO®-Spring, the treating therapist recorded the range of motion twice a week in order to individually adjust the size of the virtual playing field in the game. Both the movement tasks and range of motion were increased within the intervention time according to the pre-specified protocol [12].

Figure 1: Training on the ARMEO®Spring



Following a preceding power analysis based on the results of a non-randomised pilot study, 26 participants per group were required to demonstrate a 13 point difference in the Disabilities of the Arm, Shoulder and Hand Outcome Questionnaire (effect size = 0.8) with 80% power at a significance level of $\alpha = 0.05$ [6].

Because normal distribution of the data could not be confirmed with the Shapiro-Wilk test, the group comparisons of the outcome measures were carried out with the rank-based Wilcoxon test. The group comparisons of the outcome measures between the control and intervention group were carried out with the Mann-Whitney-U-Test. Centre affiliation, the amount of conventional therapy, and the severity of injury (AO classification; Association for the Study of Internal Fixation; [13]) of the participants were considered as co-variates. An intention-to-treat approach with mean by group imputation for all randomised participants with incomplete datasets was performed. In order to detect group differences in the baseline demographics in injury severity as well as in the extent of conservative therapy, mean differences were assessed using the Mann-Whitney-U-Test. The two-sample t-test was used to analyse age differences. Group differences in the dominance of the affected arm, gender, and the type of surgical treatment were calculated using the Chi-square test. All data were analysed using the statistical package for social sciences SPSS Version 19 (IBM Corporation, Armonk, NY, USA).

Results

A total of 923 patients were screened from January 2016 to June 2018 in the three different study centres of which 48 patients were included (Figure 2). No significant differences were found between groups for baseline demographics, clinical characteristics, or in the duration and frequency of conventional therapy during the intervention (Table 1). The baseline assessments for Disabilities of the Arm, Shoulder and Hand Questionnaire, Wolf Motor Function Test Orthopaedic, range of motion and grip strength were comparable. Three participants were lost during the intervention period, but not due to the intervention, which suggests a good acceptance of the robotic intervention (Figure 2). One patient was involved in a car accident (intervention group), one participant was too small to fit into the robotic device (intervention group), and a further participant developed an infection of the shoulder (control group). There was neither an adverse nor a serious adverse event related to the training on the robotic-assisted ARMEO®-Spring device nor to the assessments. Therefore, a safe implementation of the robotic intervention is possible.

Figure 2 Flowchart of study enrollment

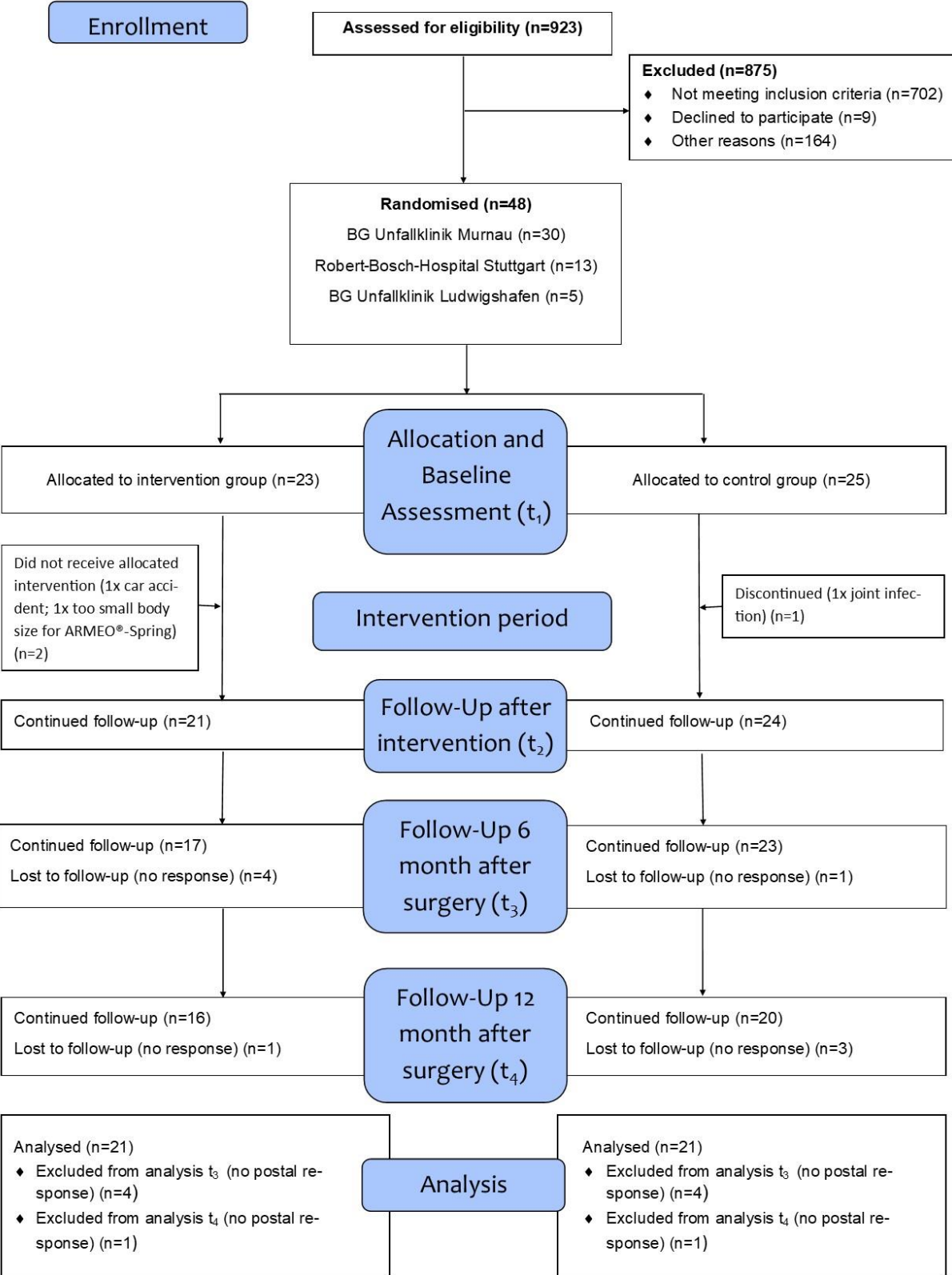


Table 1: Description of study population

Characteristic	intervention group	control group	p-value
Participants (n)	23	25	
Age (years)			
Mean (SD)	55.1 (9.8)	55.0 (9.5)	0.89
Gender			
Female (%)	15 (65%)	16 (64%)	0.93
Affected Arm ^a			0.68
Dominant arm affected (%)	11 (48%)	11 (44%)	
Non-dominant arm affected (%)	11 (48%)	14 (56%)	
AO-classification			0.45
A11-A (%)	5 (22%)	1 (4%)	
A11-B (%)	9 (39%)	14 (56%)	
A11-C (%)	9 (39%)	10 (40%)	
Surgical treatment			0.71
Plate osteosynthesis (%)	12 (52%)	17 (68%)	
Screw osteosynthesis (%)	1 (4%)	1 (4%)	
Endoprosthesis (%)	2 (9%)	1 (4%)	
Humeral nail (%)	8 (35%)	6 (24%)	
Conventional therapy (min) ^b			
Mean individual therapy (SD)	397 (461)	498 (404)	0.26
Mean group therapy (SD)	376 (527)	305 (333)	0.69

n Number, *SD* Standard deviation, *AO* Arbeitsgemeinschaft Osteosynthesefrage, *A* extraarticular, *B* partial joint fracture, *C* total joint fracture; ^aDominant arm could not be determined in one participant in the intervention group, ^b*n*= 21 intervention group, *n*= 24 control group

During the intervention period, the control group and the intervention group improved their Disabilities of the Arm, Shoulder and Hand Scores (primary endpoint) by 26% and 29%, respectively ($p \leq 0.001$) with no significant difference between the intervention and control group at any time point ($p \geq 0.46$; Table 2).

Table 2: Results of the Disabilities of the Arm, Shoulder and Hand Questionnaire for the different time points and groups

DASH [points]	Baseline	Post intervention	6 months Follow-up	13 months Follow-up
Control group (n=24)				
Mean (SD)	53 (19)	39 (18)	26 (22)	21 (21)
Δ (95%CI)		Δ^1 14 (11-18)	Δ^2 13 (7-19)	Δ^3 6 (-3-14)
p-value		<0.001	0.001	0.11
Intervention group (n=21)				
Mean (SD)	52 (21)	37 (15)	30 (19)	21 (17)
Δ (95%CI)		Δ^1 15 (8-22)	Δ^2 7 (-2-16)	Δ^3 9 (1-16)
p-value		0.001	0.11	0.01
Group difference Δ (95% CI)	1 (-12-13)	-2 (-8-12)	4 (-9-16)	1 (-11-12)
p-value	0.71	0.73	0.46	0.55

DASH Disability of Arm, Shoulder and Handy Questionnaire, Δ : mean difference, CI confidence interval, Δ^1 = baseline - post intervention, Δ^2 = post intervention - 6 month follow up, Δ^3 = 6 month follow up - 13 month follow up

Participants improved significantly in both domains of the Wolf Motor Function Test Orthopaedic during the intervention period ($p < 0.001$) with larger improvements in the quality of movement domain compared to the functional capacity domain. The control group started at baseline with lower scores for both domains. However, there was no significant difference between the control and the intervention group at follow up after intervention (Table 3, $p \geq 0.46$).

Range of motion improved in all movement categories after intervention (Table 3). The largest improvements were found for flexion ($p < 0.001$) and abduction ($p \leq 0.008$). The improvement of the range of motion was not significantly different between the intervention and control group at follow up after the intervention period ($p \geq 0.06$; Table 3), except for shoulder extension ($p = 0.04$; Table 3).

Grip strength of the affected arm was also significantly increased in both groups after the intervention period ($p \leq 0.001$). No significant difference in the improvement was found between the groups ($p = 0.29$; Table 3).

1 Table 3: Results of the Wolf Motor Function Test Orthopaedic, range of motion and grip strength for the different time points and groups

	Intervention group				Control group				Group difference*	
	Baseline <i>mean</i> (SD)	Post intervention <i>mean</i> (SD)	Δ mean (95%CI)	p-value	Baseline <i>mean</i> (SD)	Post <i>intervention</i> <i>mean</i> (SD)	Δ mean (95%CI)	p-value	mean (95%CI)	p-value
WMFT-O [points]										
Functional Capacity	56 (8)	64 (6)	8 (6-10)	<0.001	53 (11)	65 (6)	13 (9-16)	<0.001	1 (-3-5)	0.46
Quality of Movement	58 (10)	67 (6)	10 (6-14)	<0.001	52 (13)	68 (4)	16 (11-20)	<0.001	1 (-3-4)	0.50
ROM [°]										
Flexion	38 (19)	62 (21)	24 (14-32)	<0.001	40 (21)	57 (24)	17 (10-24)	<0.001	5 (-9-18)	0.25
Extension	26 (13)	29 (12)	3 (-4-10)	0.22	31 (12)	38 (13)	8 (3-13)	0.004	9 (2-17)	0.04
Abduction	46 (13)	55 (19)	10 (3-17)	0.008	39 (22)	65 (17)	26 (15-37)	<0.001	10 (-1-21)	0.06
Adduction	9 (11)	14 (15)	6 (2-10)	0.01	8 (11)	17 (17)	9 (3-15)	0.003	2 (-7-12)	0.65
Ext. Rotation	14 (15)	18 (11)	5 (0-10)	0.05	15 (13)	24 (17)	9 (3-14)	0.004	6 (-2-14)	0.26
Grip strength [kg]	25 (12)	29 (10)	4 (2-5)	<0.001	23 (9)	26 (10)	3 (2-4)	0.001	3 (-3-9)	0.29

2 * Difference between intervention group and control group after intervention, SD Standard deviation, Δ baseline – post intervention, 95% CI confidence interval,
3 WMFT-O Wolf Motor Function Test-Orthopaedic, ROM Range of motion, ° degree of range of motion, kg kilogram

Discussion

The aim of this study was to investigate the efficacy and analyse the safety and acceptability of a robot-assisted therapy added to conventional therapy over a period of three weeks in proximal humeral fracture rehabilitation. We observed a continuous functional improvement in the patient-reported outcome (Disabilities of the Arm, Shoulder and Hand Questionnaire) in both groups over a 13-month period, with shown improvement in the supervised clinical outcomes (Wolf Motor Function Test Orthopaedic, range of motion, grip strength) in both groups. In contrast to our study hypothesis, we did not observe an additional benefit of the robot-assisted therapy in the intervention group compared to the control group. The participants of the exergaming exoskeleton group had a high adherence and acceptability with no drop-out due to the treatment. The treatment on the robotic device was safe, and no adverse events were noticed.

Although the robotic therapy has not improved the outcome, participation in the robotic therapy might contribute to the motivation of the patient and to adherence to rehabilitation prescriptions. The low drop-out rate in the intervention group indicates good acceptance of ARMEO®-Spring among patients. Only two out of 23 subjects from the intervention group were excluded during the therapy phase and in both cases the exclusion of these subjects was not by patient choice. The motivating character of computer-assisted forms of therapy has been previously confirmed [22].

Patient reported outcomes are core assessments of patient satisfaction and are successfully administered for the evaluation of treatment outcomes [23]. The Disabilities of the Arm, Shoulder and Hand Questionnaire was designed to measure function of the upper extremity. One limitation of the Disabilities of the Arm, Shoulder and Hand Questionnaire with potential relevance to this study is that no distinction is made between the injured side being dominant or non-dominant [23]. The ability to perform one-handed functions in everyday life, however, is differently affected whether the dominant or non-dominant extremity is injured. From the participants included in this study, nearly a half had injured the dominant arm, and 52% had injured the non-dominant arm (in one participant dominant arm could not be determined). As this distribution was similar in both groups the respective weakness of the Disabilities of the Arm, Shoulder and Hand Questionnaire likely had no influence on the results but may have reduced the power to detect significant differences. The Disabilities of the Arm, Shoulder and Hand Questionnaire has been previously used in several studies investigating patients after proximal humeral fracture, such as Gracitelli and Davids who have compared different surgical treatments of proximal humeral fracture. Compared to our study, their findings showed slightly better or equal values of the

Disabilities of the Arm, Shoulder and Hand Questionnaire at 6 months follow-up and similar values at 12 months [24,25].

Patient reported outcomes should be complimentary to supervised clinical outcome assessments according to current standards [26, 27]. To complement the Patient reported outcomes with a supervised clinically administered outcome assessment, the Wolf Motor Function Test-Orthopaedic was used to measure functional capacity and quality of movement. The Wolf Motor Function Test Orthopaedic has previously shown a somewhat higher sensitivity to change in proximal humeral fracture patients compared to the Disabilities of the Arm, Shoulder and Hand Questionnaire [18]. Considering the results of the minimally important difference (0.3 standard deviations from baseline) of a previous study with a similar study population, it can be seen that the improvements achieved during the intervention period in the present study (Wolf Motor Function Test Orthopaedic and Disabilities of the Arm, Shoulder and Hand Questionnaire) were not only statistically significant but also clinically relevant [18].

The range of motion is a frequently used parameter to assess shoulder function. There are several possibilities to measure the range of motion. The 2-arm goniometer is the most commonly used measuring instrument for assessing range of motion [28]. In the case of active shoulder abduction in an upright position, interrater reliabilities of ICC= 0.92 in healthy [28] as well as ICC=0.69 (flexion/abduction) and ICC= 0.64 (external rotation) in shoulder injury patients [29] are reported. No conclusions could be drawn for other movement directions in terms of reliability of active measurement carried out in the upright sitting position. Due to the multi-centre structure of the present study, it was inevitable that different testers would measure the range of motion. In order to perform the measurement as homogenously as possible, each assessor received standardised training before the start of the study. Additionally, an attempt was made to ensure that the same assessor performed the assessment before and after the intervention period and therefore the assessor should have only minor effects on the reported changes over time. Studies, which examined the range of motion within the first three months after proximal humeral fracture, were not identified. In previous studies on proximal humeral fracture patients the ranges of motion have been evaluated at 6 and 12 months postoperatively [25,30] and were found to be slightly larger compared to ranges of motion in our study at the end of the intervention period at approximately 3 months.

The grip strength was measured using the Jamar dynamometer and participants were seated as recommended by the American Society of Hand Therapists [31]. A meta-analysis from Bohannon and colleagues reported grip strength reference values for males age range 50-59 years between 41 kg and

51 kg depending on the measured side and between 27 kg and 31 kg for females of the same age range [32]. According to those reference values and in view of the fact that more than 60% of the present study population consisted of women, the collected grip strength values were considered adequate at 3 months postoperatively.

The current study had some limitations. Despite a careful selection of participating study centres, the recruitment rate was slow and differed considerably among the three centres. The required sample size of 26 participants per group was not fully reached during the funding period so the study was underpowered to detect a significant difference between the two groups. As the observed difference in Disabilities of the Arm, Shoulder and Hand Questionnaire of 2 points was considerably smaller compared to the anticipated difference of 13 points, no efforts were made to extend the recruitment period. Enrolled participants received different types of surgical treatment such as plates, nails and inverse prosthesis and no patients with conservative treatment were included. A more homogeneous fracture treatment of the included participants might have reduced the variability in the outcome but would have markedly prolonged the recruitment period. As different osteosynthetic treatments were equally distributed between both groups, the different treatments probably had only minor influence on the findings. The upper age limit was set at 70 years, thus findings might differ for an older population that is also frequently affected by proximal humeral fractures. Additionally it should be noted that the control group has not received equivalent additional intervention. It may well be that robotic-assisted therapy would have performed better in verge of another additional form of therapy. Finally, as with most multicentre studies which include occupational and physiotherapeutic interventions, the individual content of the therapy was not identical for each patient. In preparation for the study we tried to make the therapeutic protocol comparable among sites and we also determined the amount of therapy obtained by each patient. Another limitation, which arose during the multicentre trial, was a systematic error with regard to the timing of the 6 and originally 12 months follow ups. As a result, there was a long period of time in the realisation of the 6 and, in particular, of the 12 month follow ups between the centres.

Although in our specific setting the additional robotic therapy has not shown to significantly improve functional outcome in proximal humeral fracture patients, it might be worthwhile to explore the adaptation of some aspects of the treatment regimen. These adaptations might include the timing and intensity of the intervention, the choice of the target group, and the content and challenge of the video games. Since the exoskeleton allows a personalised calibration, the timing of the intervention could be started earlier.

The robotic device allows that the weight support can be individually adapted by a spring mechanism. In the present study we used a conservative approach by starting the intervention between the 4th and 7th postoperative week. By individually adjusting the physical challenge of the robotic device, individual patients could start the training at an earlier point in time. As no adverse event occurred in this study and in another pilot study [6], it should be considered whether in further studies the intervention could commence earlier to allow a reduction in sick leave and a quicker return to manage activities of living. A further aspect is the duration, frequency and content of the therapy. The tested device allows training up to about 90° of flexion of the shoulder joint. Since both values for shoulder flexion in the intervention group (62°) and in the control group (57°) were still below 90° at follow up after the intervention, the intervention could have been carried out longer than 3 weeks. Whether a higher training intensity or a longer intervention period would have changed the results remains unclear and should be investigated in further studies. Furthermore, future studies could analyse the potential of substituting conventional therapy by robotic devices and use robotics for more rapid rehabilitation regimes

In conclusion, the findings from this study suggest that under the tested conditions a supplemental robotic rehabilitation does not improve the outcome in patients with fractures of the proximal humerus. The acceptance and safety profile of the robotic device might support the exploration of alternative protocols for robotic-assisted rehabilitation in patients with shoulder injuries.

Clinical Messages

- no additional improvement of the shoulder function after robot-assisted training compared to control group was found over an period of 13 month
- based on the exploratory analysis of the dropout rate and adverse event rate good acceptance and good safety of the robot-assisted gamified therapy could be assumed

Acknowledgements: The authors gratefully acknowledge K. Gordt, for their support in data collection.

Author contributions: IK, CN, LS, JAM, SSF, CB and PA made a substantial contribution to the concept of the work and in writing of the manuscript. CN, LS, SS, IK, JAM, SSF and PCN helped with the recruitment and data collection. All authors revised the manuscript critically for important intellectual content, read and approved the final version to be published.

Funding support: This work was funded by the German Social Accident Insurance (DGUV) project FP-0233: “Robotic assisted training after humeral fracture – a randomised intervention study”. The funding body had no influence on the study design and the writing of the paper.

Completing interests: All Authors declare that there is no conflict of interest.

Research ethics and patient consent: Approval for the study was obtained from the ethics committee of the state medical association Rheinland-Pfalz (837.519.15 (10287)) and the ethics committee at the medical faculty of the Eberhard-Karls-University Tübingen (381/2015BO1) and are in accord with the Declaration of Helsinki. All participants gave written informed consent.

Availability of data and materials: The datasets used and/or analysed during the current study are available from the RASTA consortium or the corresponding author upon reasonable request.

References:

- [1] Court-Brown CM, Caesar B. Epidemiology of adult fractures: A review. *Injury* 2006;37:691–7.
- [2] Passaretti D, Candela V, Sessa P, Gumina S. Epidemiology of proximal humeral fractures: a detailed survey of 711 patients in a metropolitan area. *J Shoulder Elb Surg* 2017;26:2117–24.
- [3] Einsiedel T, Becker C, Däxle M, Lechner F, Kramer M, Gebhard F, et al. Frakturen der oberen Extremität beim geriatrischen Patienten – Harmlose Monoverletzung oder Ende der Selbstständigkeit? Do injuries of the upper extremity in geriatric patients end up in helplessness? A prospective study for the outcome of distal radius an. *Z Gerontol Geriatr* 2006;39:451–61.
- [4] Hodgson S. Proximal Humerus Fracture Rehabilitation. *Clin Orthop Relat Res* 2006;447:131–8.
- [5] Zech A, Hübscher M, Vogt L, Banzer W, Hänsel F, Pfeifer K. Neuromuscular training for rehabilitation of sports injuries: a systematic review. *Med Sci Sports Exerc.* 2009;41(10):1831–41.
- [6] Schwickert L, Klenk J, Stähler A, Becker C, Lindemann U. Roboterassistierte Rehabilitation proximaler Humerusfrakturen im virtuellen Raum. Eine Pilotstudie. *Z Gerontol Geriatr* 2011;44:387–92.
- [7] Macintyre NJ, Kwan LL, Johal H, Lefavre KA, Sprague S, Bhandari M, et al. SM Gr up SM Journal of Trauma Care & Rehabilitation of Proximal Humerus Fractures- A Scoping Review 2017;1:1–12.
- [8] Veerbeek JM, Langbroek-Amersfoort AC, Van Wegen EEH, Meskers CGM, Kwakkel G. Effects of Robot-Assisted Therapy for the Upper Limb after Stroke. *Neurorehabil Neural Repair* 2017;31:107–21.
- [9] Gijbels D, Lamers I, Kerkhofs L, Alders G, Knippenberg E, Feys P. The Armeo Spring as training tool to improve upper limb functionality in multiple sclerosis: A pilot study. *J Neuroeng Rehabil* 2011;8:5.
- [10] Padilla-Castañeda MA, Sotgiu E, Barsotti M, Frisoli A, Orsini P, Martiradonna A, et al. An Orthopaedic Robotic-Assisted Rehabilitation Method of the Forearm in Virtual Reality Physiotherapy. *J Healthc Eng* 2018;2018:1–20.
- [11] Handoll HH, Brorson S. Interventions for treating proximal humeral fractures in adults. *Cochrane Database Syst Rev* 2015;2015.
- [12] Nerz C, Schwickert L, Becker C, Studier-Fischer S, Müßig JA, Augat P. Effectiveness of robot-assisted training added to conventional rehabilitation in patients with humeral fracture early after

- surgical treatment: protocol of a randomised, controlled, multicentre trial. *Trials* 2017;18:589.
- [13] Baierlein SA, Herausgeber. *Frakturklassifikationen* [Internet]. Stuttgart: Georg Thieme Verlag; 2011 [zitiert 6. Mai 2020]. Verfügbar unter: <http://www.thieme-connect.de/products/ebooks/book/10.1055/b-002-41857>
- [14] Katzman R, Brown T, Fuld P, Peck A, Schechter R, Schimmel H. Validation of a short Orientation-Memory-Concentration Test of cognitive impairment. *Am J Psychiatry*. Juni 1983;140(6):734–9.
- [15] Freyd M. The Graphic Rating Scale. *Journal of Educational Psychology*. 1923;14(2):83–102.
- [16] Oberle C, Becker C, Schölch S, Lenz JU, Studier-Fischer S, Augat P, et al. Inter-rater and intra-rater reliability of an adapted Wolf motor function test for older patients with shoulder injuries. *Z Gerontol Geriatr* 2018;51:293–300.
- [17] Wolf SL, Catlin PA, Ellis M, Archer AL, Morgan B, Piacentino A. Assessing Wolf Motor Function Test as outcome measure for research in patients after stroke. *Stroke*. 2001;32(7):1635–9.
- [18] Nerz C, Schwickert L, Schölch S, Gordt K, Nolte PC, Kröger I, Augat P, Becker C. Inter-rater reliability, sensitivity to change and responsiveness of the orthopaedic Wolf-Motor-Function-Test as functional capacity measure before and after rehabilitation in patients with proximal humeral fractures. *BMC Musculoskeletal Disorders* 2019; 20:315.
- [19] Magosch P, Lichtenberg S, Loew M, Tauber M, Habermeyer P. Die Klinische Untersuchung der Schulter. *Dtsch Z Sportmed* [Internet]. 1. Dezember 2013 [zitiert 8. April 2020];2013(12). Verfügbar unter: <http://www.zeitschrift-sportmedizin.de/artikel-online/archiv-2013/heft-12/die-klinische-untersuchung-der-schulter/>
- [20] American Society of Hand Therapists. *Clinical assessment recommendations* [Internet]. 2nd ed. Chicago: The Society; 1992 [zitiert 6. Mai 2020]. Verfügbar unter: <https://www.worldcat.org/title/clinical-assessment-recommendations/oclc/32415221>
- [21] Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, et al. The REDCap consortium: Building an international community of software partners. *J Biomed Inform*. 2019; 95.
- [22] Dahl-Popolizio S, Loman J, Cordes CC. Comparing Outcomes of Kinect Videogame-Based Occupational/Physical Therapy Versus Usual Care. *Games Health J* 2014;3:157–61.
- [23] Wylie JD. Functional outcomes assessment in shoulder surgery. *World J Orthop* 2014;5:623.
- [24] Gracitelli MEC, Malavolta EA, Assunção JH, Kojima KE, dos Reis PR, Silva JS, et al. Locking intramedullary nails compared with locking plates for two- and three-part proximal humeral

- surgical neck fractures: A randomized controlled trial. *J Shoulder Elb Surg* 2016;25:695–703.
- [25] Davids S, Allen D, Desarno M, Endres NK, Bartlett C, Shafritz A. Comparison of Locked Plating of Varus Displaced Proximal Humeral Fractures with and without Fibula Allograft Augmentation. *J Orthop Trauma* 2019.
- [26] Vincent JI, MacDermid JC, King GJ, Grewal R. Validity and sensitivity to change of patient-reported pain and disability measures for elbow pathologies. *J Orthop Sports Phys Ther.* April 2013;43(4):263–74.
- [27] Westphal T. [Reliability and responsiveness of the German version of the Disabilities of the Arm, Shoulder and Hand questionnaire (DASH)]. *Unfallchirurg.* Juni 2007;110(6):548–52.
- [28] Kolber MJ, Fuller C, Marshall J, Wright A, Hanney WJ. The reliability and concurrent validity of scapular plane shoulder elevation measurements using a digital inclinometer and goniometer. *Physiother Theory Pract* 2012;28:161–8.
- [29] Hayes K, Walton JR, Szomor ZL, Murrell GA. Reliability of five methods for assessing shoulder range of motion. *Aust J Physiother* 2014;47:289–94.
- [30] Ge W, Sun Q, Li G, Lu G, Cai M, Li S-H. Efficacy comparison of intramedullary nails, locking plates and conservative treatment for displaced proximal humeral fractures in the elderly. *Clin Interv Aging* 2017;Volume 12:2047–54.
- [31] Fess EE. Grip strength. In: Casanova JS, editor. *Clinical assessment recommendations*. 2nd ed. Chicago: American Society of Hand Therapists; 1992. p. 41–5.
- [32] Bohannon, R. W., Peolsson, A., Massy-Westropp, N., Desrosiers, J. & Bear-Lehman, J. (2006). Reference values for adult grip strength measured with a Jamar dynamometer: a descriptive meta-analysis. *Physiotherapy*, 92, 11-15.