

DASH-Assessment und Likert-Skala: ein zielgerichtetes methodisches Vorgehen zur Erhebung des MCID?

Doris Unger

Institut für Physikalische Medizin und Rehabilitation, Hanusch Krankenhaus, Wien

EINLEITUNG: In der Ergotherapie wird es immer wichtiger, Therapieerfolge messbar und dadurch nach außen sichtbar zu machen. Dafür werden Assessments zur Evaluierung eingesetzt. Um ein geeignetes Instrument auszuwählen, sollten die Gütekriterien wie Validität, Reliabilität und Veränderungsempfindlichkeit bekannt sein bzw. untersucht werden. Die Veränderungsempfindlichkeit besagt, dass auch kleine Unterschiede gemessen werden können.

ZIEL: Es soll in einer Pilotstudie untersucht werden, ob es mit Hilfe der ankerbasierten Methode möglich ist, den MCID (kleinster klinisch relevanter Unterschied) im DASH-Fragebogen (The Disabilities of the Arm, Shoulder and Hand outcome measure) für Erwachsene mit Handverletzungen distal der Handwurzel zu erheben. Der MCID ist ein Punkteunterschied, der für die getestete Person wichtig und bedeutend ist. Dieser ist wichtig, damit ErgotherapeutInnen Testergebnisse richtig interpretieren können und somit den DASH-Fragebogen zielgerichtet einsetzen können.

METHODE: PatientInnen, die Handverletzungen distal der Handwurzel erlitten hatten, wurden in die Pilotstudie eingeschlossen. Sie wurden gebeten, den DASH-Fragebogen zu drei verschiedenen Zeitpunkten auszufüllen: vor Beginn der Ergotherapie (ET), nach 6x Ergotherapie und nach 12x Ergotherapie. Zusätzlich gaben die PatientInnen und Ergotherapeutinnen nach 6x und nach 12x Therapie ihre Einschätzung bezüglich der Veränderung des Gesundheitszustandes ab. Verwendet wurde eine Likert-Skala (5 Einschätzungsmöglichkeiten). Die ankerbasierte Methode fand somit Anwendung.

ERGEBNIS: Personen nahmen an der Pilotstudie teil. Eine Person brach die Studie ab, sodass die Ergebnisse von 9 Personen verwendet werden konnten. Veränderungen von 11,4 (Median) bewerteten PatientInnen als "stark verbessert". Veränderungen von 9,2 (Median) bewerteten Therapeutinnen als "stark verbessert".

Der DASH-Fragebogen und die Likert-Skala stimmten nicht miteinander überein. Immer erfolgte die Einschätzung "verbessert" oder "stark verbessert", welche allerdings nicht von den DASH-Werten belegt werden konnte.

DISKUSSION: Die Studie wies Mängel auf. Beispiele dafür sind, dass Untersucherin und behandelnde Ergotherapeutin ein und dieselbe Person waren, oder die Datenmenge zu gering war. Die Werte (11,4 und 9,2) für "stark verbessert" kommen einem MCID noch am nächsten. Allerdings können diese Werte aufgrund der geringen Datenmenge nicht schlüssig interpretiert werden. Der MCID müsste mit Hilfe einer adaptierten Methode in einer neuen Studie erhoben werden.

SCHLÜSSELWÖRTER: Ergotherapie, DASH, MCID, ankerbasierte Methode, Handverletzung

DASH-ASSESSMENT AND LIKERT-SCALE: A DIRECT METHODOLOGY TO DETERMINE THE MCID

INTRODUCTION: In occupational therapy, it is increasingly important to measure treatment success and thereby make it visible to health care providers. Therefore, assessments are used for evaluation. To select an appropriate instrument, the quality criteria such as validity, reliability and sensitivity to change should be available or must be investigated if necessary. Sensitivity to change means that even small differences can be measured.

OBJECTIVE: A pilot study should investigate whether it is possible to define the MCID (minimum clinically important difference) of the DASH questionnaire (The Disabilities of the Arm, Shoulder and Hand outcome measure) for adults with hand injuries located distal of the wrist based on an anchor-based method. The MCID is a difference, which is important and significant for the patient. The observed difference is important for occupational therapists to interpret test results properly, and thus apply the DASH questionnaire purposefully.

METHOD: Patients who had hand injuries located distal of the wrist have been included in the pilot study. They have been asked to complete the DASH questionnaire on three occasions: before the beginning of occupational therapy (OT), after 6 sessions of occupational therapy and after 12 sessions of occupational therapy. In addition, the patients and the occupational therapists estimated after 6 sessions of OT and after 12 sessions of OT how the state of health changed. A likert scale (5 answers were possible) was used.

RESULTS: 10 people took part in the pilot study. One person dropped out of the study, so the results could be used from 9 people. The patients rated a change of 11.4 (median) as "much improved", as therapists rated a change of 9.2 (median) as "much improved".

The DASH questionnaire and the likert scale did not correspond with each other. All patients stated their condition as "improved" or "much improved"; however this could not be confirmed by the DASH-values.

DISCUSSION: The study had several flaws. Examples of these are that investigator and treating occupational therapist was one and the same person or the amount of data was too low. The values (11.4 and 9.2) for "much improved" come close to a MCID. However, these values cannot be interpreted properly because of the small amount of data. The MCID should be determined using an adapted method in a new study.

KEY WORDS: occupational therapy, DASH, MCID, anchor based method, hand injury

Einleitung

In der Ergotherapie ist bei der klinischen Erfassung handgeschädigter PatientInnen der Einsatz von Messinstrumenten (Assessments) zu Beginn der Therapie, zur Verlaufskontrolle und am Ende der Therapie nicht mehr wegzudenken. Eine differenzierte Erfassung ist richtungsweisend für die folgenden ergotherapeutischen Maßnahmen und die Behandlung allgemein (1).

Messinstrumente sollten standardisiert sein und bestimmte Kriterien (Reliabilität, Validität, Änderungssensitivität) erfüllen. Welche Interpretation lässt das Scoring-Ergebnis (Punktwert) eines Tests zu, und sind überhaupt Veränderungen über einen gewissen Zeitraum messbar - die Antworten auf diese beiden Fragen sind, wenn es um den Scoring-Prozess geht, essentiell (2).

Im Physikalischen Institut für Medizin und Rehabilitation des Hanusch Krankenhauses in Wien wird der DASH-Fragebogen als Messinstrument für Personen mit Verletzungen der oberen Extremität verwendet. DASH steht für "The Disabilities of the Arm, Shoulder and Hand outcome measure" und wird auch nachfolgend herangezogen. Der Fragebogen bezieht sich auf Alltagsaktivitäten, Körperfunktionen und die Teilhabe des Menschen.

In dieser Pilotstudie wurde geprüft, ob mit der ankerbasierten Methode unter Verwendung des DASH-Fragebogens der MCID (minimum clinically important difference, kleinster klinisch relevanter Unterschied) für Erwachsene mit Bewegungseinschränkungen nach Sehnenverletzungen und nach knöchernen Verletzungen distal der Handwurzel erhoben werden kann.

Jaeschke et al. definierten 1989 den Begriff des kleinsten klinisch relevanten Unterschieds als den kleinsten Punkteunterschied, der von PatientInnen als förderlich und bedeutungsvoll wahrgenommen werden kann (3).

Der MCID ist eine wichtige Information zur Interpretation von Scores in Fragebögen.

Es wurde zwei Forschungsfragen nachgegangen:

- 1. Kann mit der ankerbasierten Methode unter Einbeziehung des DASH-Fragebogens der MCID für Erwachsene mit Verletzungen der Hand erhoben werden?
- 2. Wenn ja, wie hoch ist dieser Wert?

Um diese beiden Fragen zu beantworten wurde eine Pilotstudie im Hanusch Krankenhaus unter Verwendung des DASH-Fragebogens und einer Likert-Skala (5-stufige Einschätzungsskala) durchgeführt.

Der kleinste klinisch relevante Unterschied des DASH-Fragebogens wurde schon für verschiedene Krankheitsbilder erhoben: Beaton et al. schlossen Schulter- und HandpatientInnen in ihre Studie ein (4). Gummesson et al. konzentrierten sich auf PatientInnen nach Akromionplastik und nach Spaltung des Karpalkanals (5).

Bisher gibt es aber nach umfangreicher Literaturrecherche in den Datenbanken PubMed und CINAHL keine Studie, in der dieser Punktwert für Erwachsene mit Handverletzungen distal der Handwurzel erhoben wurde.

Zweck der Pilotstudie - ergotherapeutische Relevanz

Da der Fragebogen alltagsrelevante Aktivitäten abfragt und sich auch mit Fragen bezüglich der Teilhabe des Menschen beschäftigt, ist der Einsatz dieses Messinstrumentes für ErgotherapeutInnen von Relevanz und Nutzen.

Der DASH-Fragebogen kann zielgerichtet eingesetzt werden, wenn Scores richtig interpretiert werden.

Die Erkenntnisse dieser Pilotstudie, die zielführende Methode zu erkennen und den MCID zu erheben, sind wichtig für AnwenderInnen wie zB ÄrztInnen oder TherapeutInnen, um selbst eine Studie durchzuführen bzw. Scoring-Ergebnisse richtig interpretieren zu können. Außerdem kann erst durch richtiges Evaluieren und Interpretieren ein für die PatientInnen bedeutungsvoller Fortschritt gemessen werden.

Der DASH-Fragebogen ist ein Selbsteinschätzungsbogen, der von PatientInnen selbst ausgefüllt wird. Er bezieht sich auf die obere Extremität. Ein Auszug davon wird in Abb. 1 gezeigt

Er wurde von Hudak et al. 1996 in Kanada entwickelt (6). Dr. Dorcas E. Beaton hat als Ergotherapeutin daran mitgearbeitet. Es bestand die Idee, nicht nur spezifische Instrumente für ein Gelenk (zB Schulter) oder für eine Diagnose (zB Karpaltunnelsyndrom) zu verwenden, sondern ein Instrument zu entwickeln, welches für die obere Extremität und für verschiedene Diagnosen verwendet werden kann. Die obere Extremität soll als funktionelle Einheit gesehen werden.

Alle Informationen über den DASH-Fragebogen können der homepage www.dash.iwh.on.ca entnommen werden. Von Solway et al. wurde ein Manual herausgegeben (7).

Der Bogen umfasst 30 Fragen, welche sich mit Körperfunktionen, Aktivitäten und der Teilhabe des Menschen mit Muskel- und Skeletterkrankungen der oberen Extremität beschäftigen.

Die getesteten Personen werden gebeten, alle Fragen gemäß ihrem Zustand in der vergangenen Woche zu beantworten. Wurden die angeführten Tätigkeiten in der vergangenen Woche nicht durchgeführt, dann soll eine Antwort ausgewählt werden, die am ehesten zutreffen würde.

Beispiele für im Bogen abgefragte Aktivitäten sind:

- ein neues oder festverschlossenes Glas öffnen
- eine Mahlzeit zubereiten
- Garten- oder Hofarbeit
- einen Pullover anziehen
- mit Fortbewegungsmitteln zu Recht zukommen

Es werden auch Symptome/Beschwerden angesprochen. Die Ausprägung von Beschwerden soll festgestellt werden. Die Fragen beziehen sich auf Schmerzen, Kribbeln, Schwächegefühl und Steifheit in Schulter, Arm oder Hand.

Die Skalierung reicht immer von 1 bis 5. Bei den Fragen 1 bis 21 (Aktivitäten) bedeutet 1 "keine Schwierigkeiten", 5 "nicht möglich". Bei der Frage 22 (Beeinträchtigung der sozialen Aktivitäten) steht 1 für "überhaupt nicht", 5 für "sehr". Bei der 23. Frage (Einschränkung in der Arbeit oder bei anderen alltäglichen Aktivitäten) reicht die Skalierung von 1 "überhaupt nicht eingeschränkt" bis 5 "nicht möglich". Die Fragen 24 bis 28 beschäftigen sich mit Schmerzen, Kribbeln, usw., wobei 1 "keine" und 5 "sehr starke" bedeutet.

Die Frage 29 (Schlafstörungen aufgrund von Schmerzen) bietet Antwortmöglichkeiten von 1 "keine Schwierigkeiten" bis 5 "nicht möglich".

Die letzte Frage geht auf Einschränkungen des Selbstvertrauens ein, hier steht 1 für "stimme überhaupt nicht zu", und 5 steht für "stimme sehr zu".

Der DASH-Score (DASH-Wert) ist eine Zahl zwischen 0 und 100. Desto höher der Score ist, desto höher ist die Beeinträchtigung oder Einschränkung in der Funktion.

Die Gütekriterien Validität, Test-Retest-Reliabilität und Veränderungsempfindlichkeit wurden für die deutsche Version des DASH-Fragebogens wie folgt positiv belegt:

Validität

2003 führten Germann et al. die Standardisierung und Validierung der deutschen DASH-Version 2.0 durch, inklusive kultureller Anpassung und konnten eine hohe Validität feststellen (8). Zu diesem Zeitpunkt hatte der Fragebogen noch 21 Fragen bezüglich der Funktion und der Aktivitäten des Alltags und 5 Fragen bezüglich der Symptome.

Der Bogen wurde an 6 verschiedenen PatientInnengruppen, die sich unterschiedlichen operativen Eingriffen an der Hand unterzogen, getestet (n=342). Die PatientInnen waren in Ludwigshafen in ergotherapeutischer und physiotherapeutischer Behandlung.

Die AutorInnen stellten die Kriteriumsvalidität fest, indem der Disability-Index (in 7 verschiedenen Lebensbereichen gibt die befragte Person den Grad der Behinderung an) als Referenzinstrument herangezogen wurde. Die Korrelation wurde mit $r=0,73$ (DASH-Funktion vgl. mit Disability-index) und mit $r=0,79$ (DASH-Symptom vgl. mit Disability-Index) bestätigt.

Bei der "Kriteriumsvalidität" geht es um den Zusammenhang zwischen den empirisch gemessenen Ergebnissen des Messinstrumentes und einem anders gemessenen empirischen ("externen") Kriterium (9).

Auch die Inhaltsvalidität konnte bestätigt werden. Es gab eine hohe Korrelation zwischen den Dimensionen Funktion und Symptom.

Bei der "Inhaltsvalidität" sollen alle Aspekte der Dimension, die gemessen wird, berücksichtigt werden. Jeder Aspekt des theoretischen Begriffs muss in den Operationalisierungen berücksichtigt werden (9).

Test-Retest-Reliabilität

Auf eine hohe Reliabilität und ebenfalls hohe Veränderungsempfindlichkeit gingen Westphal et al. ein (10). Sie konnten somit die noch ausstehenden Qualitätskriterien der deutschen DASH-Version ergänzen.

In dieser Studie wurden 71 PatientInnen mit distaler Radiusfraktur über ein Jahr beobachtet und zu drei Zeitpunkten (nach 12 Wochen, nach 13 Wochen, nach 1 Jahr) getestet.

Mittels verschiedener statistischer Berechnungen konnte eine ausgezeichnete Test-Retest-Reliabilität festgestellt werden.

DASH-FRAGEBOGEN (Disabilities of Arm, Shoulder and Hand Instrument)

Name:	Datum:
--------------	---------------

Sehr geehrte(r) PatientIn!

Dieser Fragebogen beschäftigt sich sowohl mit Ihren Beschwerden als auch mit Ihren Fähigkeiten, bestimmte Tätigkeiten auszuführen.

Bitte beantworten Sie alle Fragen gemäß Ihrem Zustand **in der vergangenen Woche**, indem Sie einfach die entsprechende Zahl ankreuzen.

Wenn Sie in der vergangenen Woche keine Gelegenheit gehabt haben, eine der unten aufgeführten Tätigkeiten durchzuführen, so wählen Sie die Antwort aus, die Ihrer Meinung nach am ehesten zutreffen würde.

Es ist nicht entscheidend, mit welchem Arm oder welcher Hand Sie diese Tätigkeiten ausüben.

Antworten Sie Ihrer Fähigkeit entsprechend, ungeachtet, wie Sie die Aufgaben durchführen konnten.

Bitte schätzen Sie Ihre Fähigkeit ein, wie Sie folgende Tätigkeiten in der vergangenen Woche durchgeführt haben, indem Sie die entsprechende Zahl ankreuzen.

		Keine Schwierigkeiten	Geringe Schwierigkeiten	Mäßige Schwierigkeiten	Erhebliche Schwierigkeiten	Nicht möglich
1.	Ein neues oder festverschlossenes Glas öffnen	1	2	3	4	5
2.	Schreiben	1	2	3	4	5
3.	Einen Schlüssel umdrehen	1	2	3	4	5
4.	Eine Mahlzeit zubereiten	1	2	3	4	5
5.	Eine schwere Tür aufstoßen	1	2	3	4	5
6.	Einen Gegenstand über Kopfhöhe auf ein Regal stellen	1	2	3	4	5
7.	Schwere Hausarbeit (zB Wände abwaschen, Boden putzen)	1	2	3	4	5
8.	Garten- oder Hofarbeit	1	2	3	4	5
9.	Betten machen	1	2	3	4	5
10.	Eine Einkaufstasche oder einen Aktenkoffer tragen	1	2	3	4	5

Abbildung 1
Auszug aus dem DASH - Fragebogen

Veränderungsempfindlichkeit

Der Unterschied zwischen den Werten nach 12 Wochen und nach 1 Jahr waren statistisch signifikant ($p < 0,001$). Somit konnte eine gute Veränderungsempfindlichkeit nachgewiesen werden. Die mittlere Differenz betrug 9 Punkte.

Der DASH-Fragebogen wurde daher als geeignetes Testinstrument für diese Pilotstudie ausgewählt, da er alle notwendigen Gütekriterien erfüllt.

Methodik

Die Ethikkommission der Stadt Wien wurde von dieser Pilotstudie in Kenntnis gesetzt. Die Studie wurde genehmigt. Die TeilnehmerInnen wurden zu Beginn über die Studie aufgeklärt. Wenn diese dann ihr Einverständnis gegeben haben, füllten sie eine Einverständniserklärung zur Datenerhebung und Datennutzung aus.

Um den kleinsten klinisch relevanten Unterschied des DASH-Fragebogens bei Erwachsenen mit Verletzungen der Hand zu erheben, wurde folgendes methodische Vorgehen umgesetzt: dieses Testinstrument wurde vor Beginn der ergotherapeutischen Behandlung, nach 6 Therapien und nach 12 Therapien zur Erhebung herangezogen.

Danach konnten sogenannte Change Scores berechnet werden. Das ist die Punktzahl, um die sich zwei Erhebungen unterscheiden.

Zusätzlich wurde eine 5-stufige Likert-Skala zur Erhebung des subjektiven Eindrucks der PatientInnen und der Ergotherapeutinnen (zwei Ergotherapeutinnen testeten die PatientInnen) bezüglich der Veränderung des Gesundheitszustandes (Verbesserung, Verschlechterung) verwendet. Diese Skala kam nach 6x Therapie und nach 12x Therapie zum Einsatz. Jaeschke et al. beschrieben die Anwendung einer mehrstufigen (7-stufigen) Likert-Skala und empfehlen diese in der klinischen Praxis (11).

Die Likert-Skala diente als Ankerskala. Revicki et al. beschrieben in einem Literaturreview die ankerbasierte Methode im Zusammenhang mit der Evaluierung des kleinsten klinisch relevanten Unterschiedes bei PatientInnenfragebögen (12).

Es konnte festgestellt werden, dass bei der ankerbasierten Methode mehrere Anker (BeurteilerInnen) empfohlen werden (zB KlinikerInnen, PatientInnen). Diese Personen geben ihre Meinung bezüglich der Veränderung des Gesundheitszustandes zwischen zwei Testungen ab.

Auch in dieser Pilotstudie wurden zwei Anker herangezogen: die PatientInnen und die Ergotherapeutinnen.

Die Einschätzung erfolgte so, dass die PatientInnen den folgenden Satz vervollständigen mussten: "Mein Gesundheitszustand hat sich durch die ergotherapeutische Behandlung ...". Dann wurde eines der fünf möglichen Wörter eingefügt: stark verbessert, verbessert, nicht verändert, verschlechtert oder stark verschlechtert.

TeilnehmerInnen der Pilotstudie

Für die Erhebung wurden PatientInnen ausgewählt, die ambulant im Hanusch Krankenhaus Ergotherapie erhielten und die Einschlusskriterien erfüllten.

Einschlusskriterien

Es handelte sich um Erwachsene (Alter ab 18 Jahren) mit Bewegungseinschränkungen nach Sehnenverletzungen und nach knöchernen Verletzungen distal der Handwurzel. Die Verletzungen waren maximal vier Monate alt.

In dieser Studie wurde der Mindestscore bei der ersten DASH-Erhebung mit 25 festgelegt, damit PatientInnen eine eventuelle Verbesserung auch noch wahrnehmen können.

Ausschlusskriterien

PatientInnen mit zusätzlichen Verletzungen an der Handwurzel, Unterarm, Ellbogen, Oberarm oder Schulter wurden ausgeschlossen. Jene, die den Fragebogen nicht alleine ausfüllen konnten (zB aus kognitiven Gründen, aufgrund mangelnder Deutschkenntnisse, aufgrund mangelnder Compliance) wurden ausgeschlossen.

Für dieses Projekt wurde eine Fallzahl von 25 Personen auf folgende Weise geschätzt: Es wurden retrospektiv die Krankenakten von Jänner bis April 2010 durchgesehen. Anschließend erfolgte eine Beurteilung der Veränderung des Gesundheitszustandes der PatientInnen (Verbesserung, Verschlechterung) durch die Ergotherapeutin. Es galt die 5 möglichen Bereiche der Likert-Skala zu füllen und die Häufigkeiten in den Bereichen zu eruieren. Nach 23 gegebenen Antworten waren 3 Bereiche mit mindestens 5 Antworten gefüllt, 2 blieben leer: 48 % (11 Personen) waren "Stark verbessert", 30 % (7 Personen) zeigten sich verbessert, 22 % (5 Personen) waren nicht verändert. Für verschlechtert und stark verschlechtert gab es keine Angaben (jeweils 0 %).

Statistische Auswertung

Alle Berechnungen erfolgten im Softwareprogramm Excel. Es wurden Change Scores berechnet. Ausge-

Tabelle. 1
Demographische Daten der TeilnehmerInnen der Pilotstudie

TeilnehmerInnen der Pilotstudie n=9	
Alter:	
Mittelwert ± Standardabweichung(Bereich)	49 ±14 (30-72) Jahre
Geschlecht:	
männlich	7
weiblich	2
betroffene Seite:	
rechts	4
links	5
dominante Seite:	
links	0
rechts	9

hend davon wurden Median-Werte und Konfidenzintervalle berechnet.

Um den durchschnittlichen Baseline Score (also den durchschnittlichen Wert der ersten DASH-Erhebung) zu erheben, wurde der Median und das 99 % - Vertrauensintervall berechnet. Mittelwerte wurden aufgrund der Ausreißer und der zu starken Verzerrung der Ergebnisse nicht berechnet. Das Vertrauensintervall wurde mit 99 % festgelegt, da es sich lediglich um eine kleine Datenmenge handelt, und dann das Vertrauensintervall größtmöglich sein soll.

Ergebnisse

Die Studie wurde über ein Jahr durchgeführt. Es nahmen 10 Personen teil. Eine Person hat den DASH-Fragebogen nur ein Mal ausgefüllt, da sie, ohne Angabe von Gründen, nicht mehr zur ergotherapeutischen

Tab elle 2:
Anzahl der ergotherapeutischen Behandlungen und Dauer in Tagen

Anzahl der Behandlungen	
Mittelwert ±Standardabweichung	10±3
Dauer in Tagen	
Mittelwert ± Standardabweichung	34 ±14

Behandlung erschien. Da der DASH- Fragebogen aber mindestens zwei Mal ausgefüllt werden musste, können hier keine Aussagen getroffen werden. Daher zählt diese Person als Einzige zur Drop-Out-Quote.

Daraus folgt, dass die Ergebnisse von 9 Personen dargestellt und analysiert werden.

Alle demographischen Daten sind in Tabelle 1 dargestellt. Es nahmen 7 Männer und 2 Frauen teil, diese

Table 3
DASH-Scores zu 3 verschiedenen Zeitpunkten und Change Scores im DASH

Person	Baseline-Scores	Scores 2. Erhebung	Scores 3. Erhebung	Change Scores 1-2	Change Scores 2-3
				erste zur zweiten Messung	zweite zur dritten Messung
1	37,8	12,5	3,3	25,3	9,2
2	52,5	39,7	37,5	12,8	2,2
3	17,5	5,8	-	11,7	-
4	31,9	30,0	-	1,9	-
5	56,3	58,3	50,9	-2	7,4
6	29,2	17,5	16,7	11,7	0,8
7	23,1	24,1	13,0	-1	11,1
8	41,5	45,8	51,7	-4,3	-5,9
9	43,3	26,7	-	16,6	-
Median	37,8	26,7	27,1	11,7	4,8
99% -Vertrauens intervall	25,9 – 48,1	14,6 – 43,2	7,1 – 50,6	-0,4 – 16,6	-2,6 – 10,8

sind im Alter von 30 bis 72 Jahren. Das durchschnittliche Alter beträgt 49 Jahre (SD 14).

Alle 9 TeilnehmerInnen sind RechtshänderInnen. Davon sind 4 auch auf der dominanten Seite verletzt gewesen. Zu den häufigsten Diagnosen zählten Frakturen, Sehnenverletzungen und Luxationen.

Durchschnittlich erhielten die Personen 10x Ergotherapie (SD 3). Der gesamte Therapiezeitraum umfasste im Durchschnitt 34 Tage (SD 14). Dies kann Tabelle 2 entnommen werden.

Die Baseline Scores (siehe Tab. 3) der ersten DASH-Fragebögen reichten von 17,5 bis 56,3 Punkte. Es wurde der Median berechnet, dieser beträgt 37,8. Das dazugehörige 99% - Vertrauensintervall [25,9 bis 48,1] lässt zwischen der unteren und oberen Wertegrenze einen großen Spielraum zu. Die Streuung der gemessenen Werte um den Mittelwert ist daher sehr groß.

Der Median der Change Scores 1-2 wurde berechnet, um den Wert zu erhalten, um den sich die PatientInnen von der ersten zur zweiten Messung verbesserten. Dieser beträgt 11,7 [-0,4 bis 16,6]. Von der zweiten zur dritten Messung verbesserten sich die PatientInnen lediglich um 4,8 Punkte [-2,6 bis 10,8]. Dies ist Tabelle 3 zu entnehmen.

Generell ist zu erwähnen, dass die Punktezahl von der ersten bis zur dritten Messung variierte. Das heißt, dass diese anstiegen und sich dann wieder verringerten oder sich nur verringerten. Nur bei einer Person stiegen die Punkte stets an, die Person selbst stufte sich allerdings immer als "verbessert" ein.

Bezüglich der Einschätzungen ist zu sagen, dass nur die Kategorien "stark verbessert", "verbessert" und "nicht verändert" vertreten sind.

Am interessantesten sind die Werte, die berechnet werden konnten, wenn die Change Scores 1-2 und 2-3 zusammengefasst wurden, da in diesem Falle der Berechnung die meisten Schätzungen zur Verfügung standen. Es konnten also die meisten Werte zur Berechnung herangezogen werden. Tabelle 4 zeigt die Ergebnisse.

Tabelle 4

Einschätzungen der Therapeutinnen und der PatientInnen zusammengefasst aus Erhebung 1-2 und 2-3: "stark verbessert" / "verbessert"

	stark verbessert	verbessert		
	PatientInnen	Therapeutinnen	PatientInnen	Therapeutinnen
Median	11,4	9,2	1,9	1,9
99 % - KI	4,8 – 23,8	6,5 – 12,0	-2,1 – 9,4	-2,4 – 12,8

Es wurden alle Werte derer PatientInnen zusammengefasst, die angaben, dass sie sich "stark verbessert" hatten. Es wurden daher die Change Scores 1-2 und 2-3 zusammengefasst. Der Median dieser Gruppe beträgt 11,4 Punkte [99% Vertrauensintervall 4,8 bis 23,8].

Auch die PatientInnengruppe, die sich "verbessert" hat, wurde auf diese Art und Weise zusammengefasst. Der Median beträgt 1,9 Punkte [99% Vertrauensintervall -2,1 bis 9,4].

Die Wertung der Therapeutinnen war punktemäßig ähnlich der Wertung der PatientInnen. Die "stark verbesserten" Personen hatten einen Median von 9,2 Punkten [99% Vertrauensintervall 6,5 bis 12,0]. Die "verbesserten" Personen hatten einen Median von 1,9 [99% -Vertrauensintervall 2,4 bis 12,8].

Der DASH-Fragebogen wurde von 6 Personen zu 3 verschiedenen Zeitpunkten und von 3 Personen zu 2 verschiedenen Zeitpunkten ausgefüllt.

Diskussion

Die Pilotstudie wies Mängel auf. Denn die Einschätzung der PatientInnen war immer "stark verbessert" oder "verbessert", egal ob die Punkte im DASH mehr wurden (die Person sich also punktemäßig verschlechterte). Daher wird nun genau analysiert, warum das DASH-Assessment und die Likert-Skala in dieser Erhebung auf keine Art und Weise übereingestimmt haben.

Einschlusskriterien

Es wurden PatientInnen in die Pilotstudie eingeschlossen, die einen niedrigeren Baseline-Score im DASH hatten als die zuvor festgelegten 25 Punkte. Diese Personen wurden eingeschlossen, damit die Fallzahl groß genug war. Das führte dazu, dass beispielsweise eine Person nur wenige Male (5x) in die Ergotherapie zur Behandlung kam, da sie bald wieder genesen war. Das heißt, dass die Möglichkeit, drei Mal zu testen, entfallen ist.

Zusätzlich ist zu erwähnen, dass der Baseline-Score generell mit 25 zu niedrig angesetzt ist, und die Patient-

Innen eventuell schon in ihren Funktionen "zu gut" waren. 5 Personen kamen höchstens 10x in Therapie. Auch hier wurde nicht immer 3x getestet.

Bei einer größer angelegten Studie wäre empfehlenswert den Baseline-Score auf mindestens 35 bis 40 Punkte anzuheben. Dann wäre es den PatientInnen eventuell noch besser möglich, Veränderungen wahrzunehmen (nämlich tatsächlich anhand von Punkten im DASH, nicht nur anhand der Einschätzungsskala).

Des Weiteren ist es möglich, dass die Arten der Verletzungen schon von vornherein "zu leicht" sind, als dass die getesteten Personen bedeutende Unterschiede wahrnehmen könnten.

Greenslade et al. belegten ebenfalls in ihrer Studie (es wurden PatientInnen mit Karpaltunnelsyndrom eingeschlossen), dass bei Personen mit niedrigeren DASH-Werten auch nur noch ein geringer Spielraum bleibt, um Verbesserungen zu erkennen (13).

Ein Einschlusskriterium war, dass Personen mit vier Monate alten Verletzungen behandelt werden. Es könnte auch möglich sein, dass 4 Monate ein zu großer Zeitraum ist, also die Verletzung schon gut verheilt ist. Hier wäre ein Zeitraum von höchstens 4 Wochen besser.

Fallzahl

Eine wesentliche Schwierigkeit bei der Durchführung der Pilotstudie war die Rekrutierung der TeilnehmerInnen. Zahlreiche PatientInnen, die dem Physikalischen Institut zugewiesen wurden, erfüllten die Einschlusskriterien nicht.

Manche Personen waren unter 18 Jahren. TeilnehmerInnen, die Handwurzelverletzungen aufwiesen (wurden oft dem Physikalischen Institut zugewiesen), konnten ebenfalls nicht teilnehmen, genauso wenig wie jene, die zwar Mittelhandverletzungen oder Finger-Verletzungen hatten, aber zusätzlich noch an beispielsweise der Schulter betroffen waren.

Zahlreiche PatientInnen hatten zwar Verletzungen in den "passenden" Körperregionen, allerdings waren sie schon so selbständig im Alltag, dass der Baseline-Score zu niedrig war, und sie somit von der Studie ausgeschlossen wurden.

Bei anderen Personen war die körperliche Einschränkung so stark, dass der Fragebogen nicht alleine ausgefüllt werden konnte.

Zu guter Letzt konnten potentielle TeilnehmerInnen den DASH-Fragebogen nicht ohne fremde Hilfe ausfüllen. Sie verstanden die schriftlichen Anweisungen

aufgrund sprachlicher oder kognitiver Probleme zu wenig. Diese Personen wurden ausgeschlossen. Generell wurden dem Institut für Physikalische Medizin und Rehabilitation zu wenige PatientInnen mit dem gewünschten Krankheitsbild zugewiesen.

Die neun PatientInnen, die an der Studie teilgenommen haben, waren nicht alle 12x in der Ergotherapie, und daher haben auch nicht alle 3x den DASH-Fragebogen ausgefüllt.

Manche Personen haben den Bogen nur 2x ausgefüllt (und waren beispielsweise nur 5x in Therapie), andere haben ihn 3x ausgefüllt, waren allerdings nur 8x oder 10x in Therapie. All diese Umstände führten dazu, dass es zu geringen verwertbaren Datenmengen gekommen ist.

UntersucherInnen, die eine nachfolgende Studie durchführen wollen, müssten eine Fallzahl von mindestens 50 in Betracht ziehen, damit die Ergebnisse aussagekräftig sind und schlüssig interpretiert werden können, und schließlich der MCID für diese PatientInnengruppe erhoben werden kann.

Optimal wäre es, wenn PatientInnen 3x den DASH ausfüllen würden, um die größtmögliche Datenmenge zu erhalten.

Die Likert-Skala

Es ist möglich, dass der verwendete Begriff "Gesundheitszustand" von der Testperson zu global verstanden wurde: unter dem Motto: *"Mir geht es allgemein gut, ich habe keine Schmerzen. Also hat sich mein Gesundheitszustand verbessert."* Der Nachsatz wäre dann: *"Aber die Hand kann ich im Alltag noch immer nicht perfekt einsetzen."* Gesundheitszustand wurde eventuell zu stark mit Gesundheit allgemein in Verbindung gebracht und von der Handverletzung separiert.

Obwohl die Person ganzheitlich gesehen werden soll, wäre eine eventuelle Adaptierung der Austausch des Begriffes "Gesundheitszustand". Dieser könnte durch "Handlungsfähigkeit im Alltag" oder "gewünschter Einsatz der Hand im Alltag" ersetzt werden. Es wäre also möglich, dass die PatientInnen das tatsächlich Gefühle nicht wahrheitsgetreu verschriftlichen konnten.

Bei der Likert-Skala besteht die Möglichkeit, dass die PatientInnen ein gewünschtes Ergebnis liefern wollten, indem sie entweder "stark verbessert" oder "verbessert" ankreuzten. In diesem Zusammenhang wird die Ergotherapeutin genannt, die hier als Störfaktor gelten kann.

In diesem Fall ist die Veränderungsempfindlichkeit der Likert-Skala fraglich. Denn die PatientInnen haben immer entweder "*stark verbessert*" oder "*verbessert*" angekreuzt. Nur ein einziges Mal hat eine Therapeutin "*nicht verändert*" als Antwort gegeben. Hier stellt sich die Frage, ob die Einschätzung der PatientInnen mit Hilfe von 5 Antwortmöglichkeiten wahrheitsgetreu erfasst werden kann, und eben Veränderungen gemessen werden können.

Die Ergotherapeutin als behandelnde Therapeutin und Untersucherin

Die Ergotherapeutin ist als Störfaktor zu nennen. Denn während der Therapie kommt es unweigerlich zu Gesprächen und auch zu Auskünften bezüglich des Wohlbefindens der PatientInnen. Es kommt zu Gesprächen bezüglich Fortschritten und Rückschlägen. Diese können eventuell PatientInnen und die Therapeutin beeinflussen.

Daher wäre bei einer nachfolgenden Studie darauf zu achten, dass die Testungen eine dritte Person übernimmt.

Zeiträume zwischen den Testungen

Die Zeiträume zwischen den Testungen waren sehr kurz. Die durchschnittliche Behandlungsdauer war 34 Tage (SD 14). Innerhalb dieser Zeit wurden zwei oder drei Testungen durchgeführt. Daraus folgt, dass eventuell keine großen wahrnehmbaren punktemäßigen Veränderungen auftreten.

Zusätzlich muss aber erwähnt werden, dass von der ersten zur zweiten Testung eine größere Veränderung als von der zweiten zur dritten Testung zu verzeichnen ist. Der Median beträgt zuerst 11,7, später nur noch 4,8 Punkte. Das kann daran liegen, dass zu Beginn noch größere Fortschritte gemacht werden als im zweiten Beobachtungszeitraum.

Bei einer größer angelegten Studie müsste darauf geachtet werden, dass mehrere Wochen (zB mindestens 5 Wochen) zwischen zwei Messungen liegen. Dann müsste natürlich auch die Anzahl der Therapien angehoben werden: beispielsweise: 10 absolvierte Therapien bei der ersten Einschätzung des Gesundheitszustandes und 20 absolvierte Therapien bei der zweiten Einschätzung des Gesundheitszustandes. Hierbei stellt sich aber dann noch die Frage, ob die getesteten Personen so viele Therapien überhaupt benötigen.

Übereinstimmung des DASH-Fragebogens mit der Likert-Skala

Am interessantesten erscheint der Zusammenhang zwischen den Ergebnissen des DASH-Fragebogens und den Ergebnissen der Likert-Skala.

In Abbildung 2 sind die Vertrauensintervalle aller Wertungen für "*stark verbessert*" und "*verbessert*" (PatientInnen und TherapeutInnen) dargestellt. Die Berechnung erfolgte aus den Change Scores 1-2. Es handelt sich dabei also um die Change Scores, die sich aus der Differenz von der ersten zur zweiten DASH-Erhebung ergaben.

Zum Zeitpunkt der zweiten DASH-Erhebung mussten die TeilnehmerInnen zum ersten Mal eine Antwort auf der Likert-Skala geben. Auch die ErgotherapeutInnen schätzten die getesteten Personen ein.

2 PatientInnen wählten "*stark verbessert*", 7 wählten "*verbessert*". Die ErgotherapeutInnen wählten 8x "*verbessert*". Eine Therapeutin entschied sich 1x für "*nicht verändert*". Es wurden in Abbildung 2 nur die Ergebnisse für "*verbessert*" und "*stark verbessert*" dargestellt.

Der Median für die Change Scores der PatientInnen in der Kategorie "*stark verbessert*" beträgt 18,5 [1,0 bis 36,0]. Der Median für die Kategorie "*verbessert*" beträgt 1,9 [-3,1 bis 13,3]. Der Median für die Change Scores der TherapeutInnen in der Kategorie "*verbessert*" beträgt 6,8 [-2,1 bis 17,0]. Keine Therapeutin hat PatientInnen als "*stark verbessert*" eingeschätzt.

In dieser Abbildung entspricht die Länge eines Balkens einem Vertrauensintervall. Drei Vertrauensintervalle sind abgebildet.

Die Vertrauensintervalle für "*verbessert*" sind bei PatientInnen und TherapeutInnen ähnlich groß. Die Einschätzung der PatientInnen als "*stark verbessert*" hat ein sehr großes Vertrauensintervall.

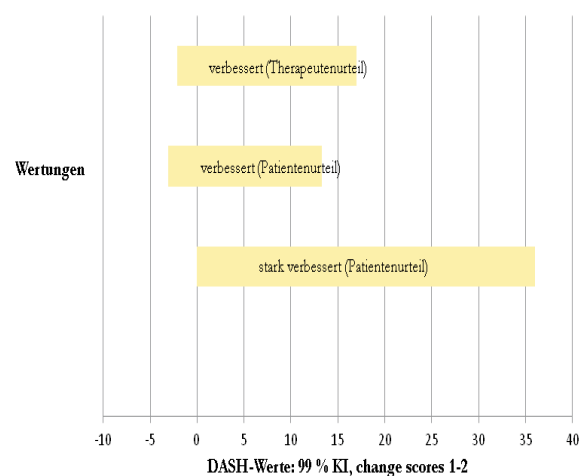


Abbildung. 2
Einschätzungen der PatientInnen und der ErgotherapeutInnen: Change Scores 1-2

In Abbildung 3 sind die Vertrauensintervalle aller Wertungen (PatientInnen und TherapeutInnen) für den 3. Erhebungszeitpunkt dargestellt. Zur Berechnung wurden die Change Scores 2-3 herangezogen. Das sind jene Change Scores, die sich aus der Differenz von der 2. zur 3. DASH-Erhebung ergeben.

2x wählten PatientInnen "stark verbessert" als Antwort, 4x wurde von ihnen "verbessert" angegeben. Die ErgotherapeutInnen schätzten die PatientInnen 3x als "stark verbessert" und 3x als "verbessert" ein.

Folgende Vertrauensintervalle sind in Form von Balken abgebildet: Der Median für die Kategorie "stark verbessert" (PatientInnen) beträgt 10,2 [7,7 bis 12,6]. Der Median für die Kategorie "verbessert" (PatientInnen) beträgt 1,5 [-5,9 bis 8,2].

Der Median für die Schätzungen der TherapeutInnen in der Kategorie "stark verbessert" ist 9,2 [6,5 bis 12]. Für "verbessert" ergibt sich ein Median von 0,8 [-7,4 bis 5,5].

In dieser Graphik fällt auf, dass sowohl die Vertrauensintervalle der Kategorie "stark verbessert" als auch die der Kategorie "verbessert" bei TherapeutInnen und PatientInnen sehr ähnlich sind. Die Schätzungen zum 3. Erhebungszeitpunkt sind daher ähnlich ausgefallen.

In Abbildung 4 sind die Vertrauensintervalle aller Wertungen dargestellt. Zur Berechnung wurden die Change Scores 1-2 und 2-3 zusammengefasst und addiert.

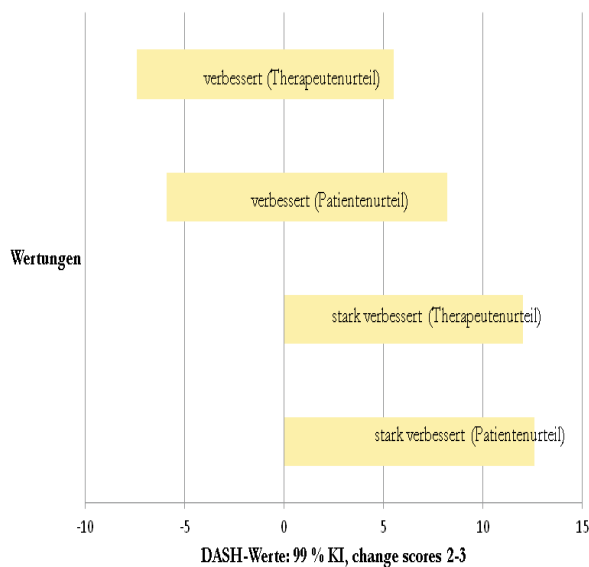


Abbildung 3
:Einschätzungen der PatientInnen und der ErgotherapeutInnen: Change Scores 2-3

4x wurde von PatientInnen "stark verbessert" angegeben, 11x "verbessert". TherapeutInnen schätzten die getesteten Personen 3x mit "stark verbessert", 11x mit "verbessert" und 1x mit "nicht verändert" ein.

Folgende Vertrauensintervalle sind graphisch dargestellt: Der Median für PatientInnen mit der Wertung "stark verbessert" beträgt 11,4 [4,8 bis 23,8]. Der Median für TherapeutInnen mit der Wertung "stark verbessert" beträgt 9,2 [6,5 bis 12,0]. Der Median für PatientInnen mit der Wertung "verbessert" ist 1,9 [-2,1 bis 9,4]. Der Median für TherapeutInnen mit der Wertung "verbessert" ist 1,9 [-2,4 bis 12,8]. Die PatientInnen vergaben ähnlich viele Punkte (Median 11,4) für "stark verbessert", wie die TherapeutInnen (Median 9,2).

In der Literatur finden sich Angaben zur Größe des kleinsten klinisch relevanten Unterschied erhoben. Beaton et al. legten einen wesentlichen Punkteunterschied mit 15 Punkten fest (4). Hingegen wurde von Gummesson et al. der kleinste klinisch relevante Unterschied im DASH mit einem Wert von 10 angegeben (5).

Daraus kann geschlussfolgert werden, dass die Werte 11,4 bzw. 9,2 dem kleinsten klinisch relevanten Unterschied noch am nächsten kommen. Trotzdem kann der MCID aufgrund der kleinen Datenmenge dieser Pilotstudie nicht berechnet werden.

Nebenbei ist noch zu erwähnen, dass das Vertrauensintervall der Einschätzungen der TherapeutInnen für "stark verbessert" kleiner war [6,5 bis 12,0] als das der

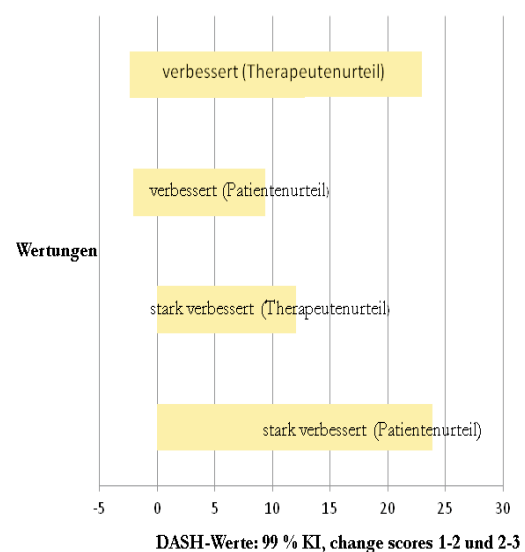


Abbildung 4
Einschätzungen der PatientInnen und der ErgotherapeutInnen: Change Scores 1-2 und 2-3

PatientInnen [4,8 bis 23,8]. Das kleinere Vertrauensintervall kann einer realistischeren Einschätzung seitens der Therapeutin entsprechen.

Die sowohl von den PatientInnen als auch von den Therapeutinnen vergebenen Werte für "verbessert" entsprechen einem Median von 1,9. Dieser ist nicht als den kleinsten klinisch relevanten Unterschied zu werten, da es gewiss nicht möglich ist, 1,9 Punkte subjektiv zu erkennen und als bedeutend einzustufen.

Des Weiteren beginnen die Vertrauensintervalle bei den PatientInnen und bei den Therapeutinnen für die Wertung "verbessert" im Minusbereich: [-2,1 bis 9,4] und [-2,4 bis 12,8]. Das bedeutet, dass eine Verschlechterung im DASH eine Verbesserung auf der Likert-Skala darstellt. Das ist paradox, und hier stimmen diese beiden Instrumente nicht überein.

Die ankerbasierte Methode ist per se eine sehr gute Möglichkeit (3); um die Einschätzung der PatientInnen zusätzlich zu einem standardisierten Test einzuholen. Diese Methode sollte daher bei einer nachfolgenden Studie beibehalten werden, allerdings sollten Änderungen mit einbezogen werden.

Schlussfolgerung

Für eine nächste Studie müsste die Methode adaptiert werden und folgende Änderungen vorgenommen werden:

- eine Fallzahl von 50
- Testerin und Ergotherapeutin sind zwei verschiedene Personen
- Baseline-Scores zwischen 35 und 40
- Die Verletzung ist höchstens 4 Wochen alt.
- Die PatientInnen müssen 3x getestet werden.
- Zwischen den Testungen liegen mindestens 5 Wochen.
- Der Begriff "*Gesundheitszustand*" wird ersetzt.

PatientInnen und Therapeutinnen schätzten Punkteunterschiede von 11,4 und 9,2 (Median) als "*stark verbessert*" ein. Diese Werte kommen einem MCID noch am nächsten. Allerdings sind diese Ergebnisse aufgrund der niedrigen Fallzahl und der geringen Datenmenge nicht schlüssig interpretierbar. Somit konnte der MCID für diese PatientInnengruppe nicht erhoben werden.

Dennoch dienen diese Ergebnisse der weiteren Forschung. Eine nächste größer angelegte Studie kann die Verbesserungsvorschläge integrieren und so die verbesserte Methode zur Erhebung des kleinsten klinisch relevanten Unterschiedes dieser PatientInnen nutzen.

Literatur

1. Diday-Nolle AP, Breier S. Klinische Erfassung handgeschädigter Patienten. In: Waldner-Nilsson, B. (Hrsg.), Ergotherapie in der Handrehabilitation. Band 1. (S.9). Berlin, Heidelberg: Springer.1997
2. Taylor MC. Evidence-based practice for occupational therapists, second edition. Oxford, UK: Blackwell Publishing Ltd.2007
3. Jaeschke R, Singer J, Guyatt GH. Measurement of health status: Ascertaining the minimal clinically important difference. *Controlled Clinical Trials* 1989, 10, 407-415.
4. Beaton DE, Katz JN, Fossel AH, Wright JG, Tarasuk V, Bombardier C. Measuring the whole or the parts? Validity, reliability and responsiveness of the Disabilities of the Arm, Shoulder and Hand outcome measure in different regions of the upper extremity. *Journal of Hand Therapy*, 2001; 14, 128-146.
5. Gummesson C, Atroshi I, Ekdahl C. The Disabilities of the Arm, Shoulder and Hand (DASH) outcome questionnaire: Longitudinal construct validity and measuring self-rated health change after surgery. *BMC Musculoskeletal Disorders*, 2003; 4, 11.
6. Hudak PL, Amadio PC, Bombardier C, the Upper Extremity Collaborative Group. Development of an upper extremity outcome measure: The DASH (Disabilities of the Arm, Shoulder and Hand). *American Journal of Industrial Medicine*, 1996; 29, 602-608.
7. Solway S, Beaton DE, McConnell S, Bombardier C. The DASH outcome measure user's manual, second edition. Toronto: Institute for Work & Health.2002
8. 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, 13-19.
9. Schnell R, Hill PB, Esser E. Methoden der empirischen Sozialforschung. München: R. Oldenbourg Verlag.1999
10. Westphal T, Piatek S, Winckler S. Reliabilität und Veränderungssensitivität der deutschen Version des Fragebogens Arm, Schulter und Hand (DASH). *Unfallchirurg*, 2007; 110, 548-552.
11. Jaeschke R, Singer J., Guyatt GH. A comparison of seven-point and visual analogue scales: Data from a randomized trial. *Controlled Clinical Trials*, 1990; 11, 43-51.
12. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *Journal of Clinical Epidemiology*, 2008; 61, 102-109.
13. Greenslade JR, Mehta RL, Belward P, Warwick DJ. DASH and Boston questionnaire assessment of carpal tunnel syndrome outcome: What is the responsiveness of an outcome questionnaire? *Journal of Hand Surgery (British and European Volume)*, 2004; 29, 159-164.

Korrespondierende Autorin:

Doris Unger, Msc

Institut für Physikalische Medizin und Rehabilitation
Hanusch Krankenhaus; Heinrich-Collin-Straße 30

1140 Wien

email: Doris.Unger@wgkk.at