

N. Smidt¹
D. A. W. M. van der Windt^{1,2}
W. J. J. Assendelft³
A. J. Mourits¹
W. L. J. M. Devillé^{1,4}
A. F. de Winter⁵
L. M. Bouter¹

Interbeobachter-Reproduzierbarkeit der Beurteilung des Schweregrads von Beschwerden, der Griffstärke und der Druckschmerzschwelle bei Patienten mit lateraler Epikondylitis

Interobserver Reproducibility of the Assessment of Severity of Complaints, Grip Strength, and Pressure Pain Threshold in Patients with Lateral Epicondylitis

Zusammenfassung

Das Ziel der Studie bestand darin, die Reproduzierbarkeit von Befunden zu Schweregrad der Beschwerden, Griffstärke und Druckschmerzschwelle bei Patienten mit lateraler Epikondylitis im Rahmen primärer medizinischer Versorgung zu bewerten.

Die Interbeobachter-Reliabilität der Beurteilung des Schweregrads von Beschwerden sowie der Griffstärke war ausgezeichnet, die der Druckschmerzschwelle jedoch nicht zuverlässig. Griffstärke und Gesamtbefund des Schweregrads der Beschwerden sind nützliche und verlässliche Maße zur Beurteilung lateraler Epikondylitis.

Schlüsselwörter

Laterale Epikondylitis · Griffstärke · Druckschmerzschwelle · Rehabilitation

Abstract

The aim of the study was to evaluate the reproducibility of the assessment of severity of complaints, grip strength, and pressure pain threshold in patients with lateral epicondylitis in primary care.

The interobserver reliability of severity of complaints and grip strength was excellent. The measurement of pressure pain threshold showed unsatisfactory reliability. Grip strength and the overall assessment of the severity of complaints are useful and reliable measures for the assessment of lateral epicondylitis.

Key words

Lateral epicondylitis · grip strength · pressure pain threshold · rehabilitation

Einleitung

Laterale Epikondylitis ist durch Schmerz an der Außenseite des Ellenbogens charakterisiert, der bei Druck auf den Epicondylus lateralis und bei Dorsalextension des Handgelenks gegen Widerstand zunimmt [13]. Wichtigster diagnostischer Befund ist die

Provokation von Schmerz am Epicondylus lateralis durch Palpation oder spezifische manuelle Tests. In der klinischen Praxis werden diese Tests auch eingesetzt, um den Status eines Patienten zu beschreiben oder die Wirksamkeit einer Behandlung zu bewerten.

Anmerkung

Die Herausgeber haben sich beim Verlag der Erstpublikation für die Leserinnen und Leser der *manuelletherapie* um eine Zweitpublikation dieses Artikels bemüht, weil sie den Inhalt für hoch relevant halten. Erstpublikation in: *Arch Phys Med Rehabil* (2002;83:1145–1150)

Institutsangaben

¹ Institute for Research in Extramural Medicine, Faculty of Medicine, VU University Medical Centre, 1081 BT Amsterdam/NL

² Dept. of General Practice, VU University Medical Centre, 1081 BT Amsterdam/NL

³ Dept. for General Practice, Division of Public Health, Academic Medical Centre, University of Amsterdam, 1081 BT Amsterdam/NL

⁴ Netherlands Institute for Health Services Research (NIVEL), Utrecht/NL

⁵ Dept. of Social Psychiatry, University of Groningen/NL

Korrespondenzadresse

Nynke Smidt, PhD · Institute for Research in Extramural Medicine, VU University Medical Center · van der Boerhorststr 7 · NL-1081 BT Amsterdam · E-mail: n.smidt.emgo@med.vu.nl

Bibliografie

Manuelle Therapie 2005; 9: 211–217 © Georg Thieme Verlag KG Stuttgart · New York
DOI 10.1055/s-2005-858886
ISSN 1433-2671

Eine kürzlich erstellte systematische Übersichtsstudie zur Wirksamkeit von Physiotherapie bei lateraler Epikondylitis (Sichtung 23 randomisierter klinischer Studien, RCT) ergab, dass in den meisten RCT Griffstärke (19/23) und Schmerz (15/23) zur Beurteilung der Wirksamkeit der Behandlung herangezogen wurden [27]. 10 von 19 RCT gaben aber nicht genau an, ob Griffstärke als *schmerzfreie* oder *maximale* Griffstärke zu verstehen ist. Nur 9 RCT beschrieben das Messprogramm, und die große Mehrheit der Studien diskutierte nicht die Reliabilität der verwendeten Methoden. Die meisten Studien beurteilten Schmerz anhand einer visuellen Analogskala oder Likert-Skala und nur 1 Studie maß die Druckschmerzschwelle [13].

Viele RCT zu muskuloskelettalen Störungen schätzen zur Beurteilung der Wirksamkeit einer Behandlung den Schweregrad der Beschwerden ein, von denen jedoch keine die Interbeobachter-Reproduzierbarkeit derartiger Einschätzungen beurteilte [17, 36]. Obwohl sich in der Literatur mehrere Berichte über Reproduzierbarkeit von Einschätzungen der Griffstärke gesunder Personen finden, wurde dieser bei Patienten mit lateraler Epikondylitis bisher wenig Aufmerksamkeit geschenkt.

Stratford et al. [31] untersuchten, mit welcher Reliabilität derselbe Beobachter bei 35 Personen mit Tennisellenbogen die Griffstärke beurteilte. Sie fanden sowohl für schmerzfreie als auch für maximale Griffstärke hohe Intrabeobachter-Reliabilitätskoeffizienten (0,97 bzw. 0,98). Angaben zur Interbeobachter-Reliabilität bezüglich Griffstärke fehlen aber immer noch. Die Autoren benutzten das Smedley-Handdynamometer, obwohl in Handtherapiekliniken häufiger das Jamar-Handdynamometer verwendet wird [31].

Mehrere Autoren [5, 7, 8, 11, 14, 20, 22] empfehlen für klinische Praxis und Forschung den Gebrauch des Druckschmerzschwellenmessers. Als zufrieden stellend wurde die Interbeobachter-Reliabilität von Messungen der Druckschmerzschwelle bei gesunden Personen [2, 6, 21] und bei Patienten mit myofaszialen Triggerpunkten [26] angegeben. Einige Studien [6, 16] zeigten jedoch, dass die Reaktionszeit der Beobachter und die unterschiedliche Geschwindigkeit des Druckanstiegs die Reliabilität beeinflussen können. Über die Reproduzierbarkeit von Messungen der Druckschmerzschwelle bei Patienten mit lateraler Epikondylitis gibt es bisher noch keine Informationen.

Gute Reproduzierbarkeit ist Voraussetzung für eine valide Messung. Außerdem werden Patienten während langfristiger Nachsorge oder einer länger währenden Behandlung oft durch verschiedene Untersucher beurteilt. Daher sollte die Interbeobachter-Reproduzierbarkeit nachweisbar und akzeptabel sein.

Angesichts der Tatsache, dass klinische Resultate wie Griffstärke und Druckschmerzschwelle in der klinischen Praxis und für die Forschung sehr wichtig sind und sehr wenig über die Reproduzierbarkeit solcher Messergebnisse bekannt ist, haben sich die Autoren dieser Studie entschlossen, die Interbeobachter-Reproduzierbarkeit von Einschätzungen des Schweregrads der Beschwerden, der Griffstärke und der Druckschmerzschwelle bei Patienten mit lateraler Epikondylitis zu bestimmen. Die Studie in Form einer RCT zur Wirksamkeit von Maßnahmen der pri-

mären Versorgung von Patienten mit lateraler Epikondylitis durchgeführt [28].

Methoden

Patienten

Die Studienteilnehmer wurden aus der Population einer in den Niederlanden durchgeführten RCT zum Vergleich der Wirksamkeit von Kortikosteroidinjektionen, Physiotherapie und einer Abwartestrategie bei lateraler Epikondylitis rekrutiert [28].

Einschlusskriterien

- Patienten mit Schmerzen an der lateralen Ellenbogenseite, die bei Druck auf den Epicondylus lateralis und bei Dorsalextension des Handgelenks gegen Widerstand zunahmen.
- Ellenbogenbeschwerden seit mehr als 6 Wochen;
- Alter zwischen 18 und 70 Jahre;
- Fähigkeit, einen Fragebogen in holländischer Sprache auszufüllen;
- Zustimmung zur Teilnahme nach gründlicher Information.

Ausschlusskriterien

- Behandlung mit Physiotherapie oder Injektion(en) während der vergangenen 6 Monate;
- Bilaterale Symptome;
- Vorliegen von Zeichen und Symptomen, die auf irgendeine andere Ursache des Schmerzes im lateralen Ellenbogen hinwiesen (z. B. Halsprobleme);
- Angeborene oder erworbene Deformitäten des Ellenbogens;
- Operative Eingriffe am Ellenbogen;
- Luxation, Sehnenrupturen oder Frakturen im Ellenbogenbereich in den vergangenen 12 Monaten;
- Systemische Störungen des muskuloskelettalen Systems;
- Neurologische Störungen;
- Kontraindikationen zur Injektion von Kortikosteroiden.

Zwischen September 1997 und Oktober 1998 nahmen insgesamt 185 Patienten an der RCT teil. Während der Nachbeobachtung (3, 6, 12, 26 bzw. 52 Wochen nach der randomisierten Zuweisung zu einer Behandlungsgruppe) wurde eine zufällige Stichprobe von 50 beteiligten Patienten eingeladen, an der Studie zur Reproduzierbarkeit von Beurteilungen teilzunehmen. Um sicherzustellen, dass die Studie Patienten mit unterschiedlich schweren Symptomen einschloss, kamen diese in jedem beliebigen Stadium der Nachbeobachtungszeit für die Beteiligung infrage. Zusätzliche Auswahlkriterien für die Reproduzierbarkeitsstudie waren die Zustimmung zur Teilnahme nach gründlicher Information, keine beidseitigen Ellenbogenprobleme, keine operativen Eingriffe am Ellenbogen und kein Gips.

Messungen

Zwischen Juli 1998 und Januar 1999 beurteilten 2 Forschungsphysiotherapeutinnen unabhängig voneinander und in zufälliger Reihenfolge jeden der 50 Patienten. Die beiden Therapeutinnen besitzen das Diplom des *Royal Dutch College of Physiotherapists* und verfügen über mehrjährige Erfahrung in klinischer Praxis und Forschung. Sie erfuhren die Vorgehensweise bei den Messungen und wurden vor der Auswahl der Patienten gründlich geschult.

In standardisierter Form nahm eine der beiden in Gegenwart der anderen die Anamnese auf. Anschließend führten beide unabhängig voneinander bei jedem Patienten eine körperliche Untersuchung von Ellenbogen und Handgelenk durch, die die Inspektion des Ellenbogens, aktive und passive Untersuchung sowie Tests von Bewegungen gegen Widerstand umfasste. Daraufhin wurde der Schweregrad der Beschwerden anhand einer 11-Punkte-Skala (0 = keine Beschwerden, 10 = sehr starke Beschwerden) eingestuft.

Mithilfe des Jamar-Handdynamometers wurden beim stehenden Patienten mit vollständig gestrecktem Ellenbogen und Schultergelenk und Radioulnargelenk in Neutral-Null-Stellung bezüglich Rotation sowohl schmerzfreie als auch maximale Griffstärke gemessen, und zwar zuerst die schmerzfreie Griffstärke des betroffenen Armes. Das Dynamometer war nach einem Standardverfahren mithilfe standardisierter Prüfgewichte geeicht (Fehler von weniger als 4% [9]).

Die Patienten sollten das Dynamometer langsam zusammenpressen, bis ein unangenehmes Gefühl eintrat. Weder die Physiotherapeutin noch der Patient konnten während des Tests die Anzeige des Dynamometers lesen. Mit einem Ruheintervall von 20 Sekunden erfolgte eine 3-malige Messung der schmerzfreien Griffstärke. Der durchschnittliche Wert (in kg) der 3 Versuche wurde berechnet und in die Analyse aufgenommen. Die schmerzfreie Griffstärke wurde immer vor der maximalen beurteilt. Zunächst fand mit einem entsprechenden Verfahren 3-mal hintereinander die Messung der maximalen Griffstärke beim nichtbetroffenen und anschließend am betroffenen Arm statt.

Die Intraobserver-Reliabilität der 3 Messungen war sowohl für schmerzfreie als auch für maximale Griffstärke ausgezeichnet (Bereich des Intraklassen-Korrelationskoeffizienten [ICC]: 0,95 – 0,97). Jeder Proband erhielt folgende verbale Instruktionen: „Sind Sie bereit? – Drücken Sie nun zusammen, so fest Sie können [...], noch mehr [...], noch mehr [...]! – Entspannen Sie sich!“

Während der Patient mit 30° abduziertem Arm, 90° flektiertem Ellenbogen und unterstütztem Unterarm, Handgelenk und Hand saß, wurde mit einem Algometer die Druckschmerzschwelle am Epicondylus lateralis gemessen. Die Physiotherapeutin palpierete leicht die betroffene gemeinsame Extensorensehne, um den empfindlichsten Bereich zu identifizieren. Anschließend setzte sie zunächst auf der nichtbetroffenen Seite den Kopf des Algometers (Hartgummispitze, Oberfläche 1 cm²) senkrecht zur gemeinsamen Extensorensehne am Epicondylus lateralis auf, bis der Patient angab, die Druck- verändere sich in eine Schmerzempfindung. Die Druckschmerzschwelle wurde auf beiden Seiten jeweils 3-mal mit einem 20-sekündigen Ruheintervall gemessen. Die Intraobserver-Reliabilität dieser Messungen war ausgezeichnet (ICC-Bereich: 0,91 – 0,96). Für die Analyse wurde der durchschnittliche Wert der 3 Versuche (in kg/cm²) berechnet. Der maximal applizierte Druck betrug 6 kg/cm [10, 27].

Nachdem Durchführung aller Messungen verließ die erste Physiotherapeutin den Raum, und die zweite Physiotherapeutin

wiederholte denselben Vorgang. Um Abfolgeeffekte zwischen den Physiotherapeutinnen zu verhindern, war die Reihenfolge beider bei der Befunderhebung gemäß dem Zeitpunkt der Befunderhebung (≤ od. > 6 Wochen nach Beginn der RCT) randomisiert und stratifiziert. Die beiden Therapeutinnen zeichneten ihre Untersuchungsergebnisse (Anamnese, körperliche Untersuchung, Gesamtbeurteilung des Schweregrades, Griffstärke, Druckschmerzschwelle) unabhängig voneinander in standardisierten Untersuchungsformularen auf. Ihre Befunde wurden erst nach Abschluss der Untersuchung bei allen 50 Patienten diskutiert. Im Rahmen der randomisierten Studie waren die Physiotherapeutinnen auch blind gegenüber der den Patienten verordneten Behandlung, damit diese Information nicht ihre Beurteilung beeinflussen konnte.

Analyse der Reproduzierbarkeit

Quantifiziert wurde die Interobserver-Reproduzierbarkeit mithilfe der Bland-und-Altman-Methode zur Einschätzung der Übereinstimmung [4] und des ICC [25] zur Bewertung der Reliabilität.

Die Bland-und-Altman-Methode gibt Aufschluss über die Verteilung der Unterschiede zwischen Beobachtern [34] und liefert Größe, Richtung und Bereich der Unterschiede zwischen Beobachtern in denselben Einheiten wie die ursprüngliche Skala. Die Interobserver-Übereinstimmung (Bland und Altman) wurde durch Berechnung der mittleren Differenz (d) zwischen den beiden Beobachterinnen und der Standardabweichung (SD) dieser Differenz quantifiziert. Je näher d an Null liegt und je kleiner die Standardabweichung dieser Differenz, umso größer die Interobserver-Übereinstimmung.

Die bezogen auf den Mittelwert der Messwerte grafisch dargestellten Unterschiede zwischen den beiden Beobachterinnen umfassen Größe, Richtung und Bereich der Unterschiede und zeigen, ob diese zwischen den Beobachtern im ganzen Messbereich konsistent sind. Die 95%-Grenzen der Übereinstimmung wurden als mittlere Differenz zwischen Beobachtern ± 1,96 SD der Differenzen definiert; dies gibt den Gesamtfehler an (systematischer plus zufälliger Fehler). Ein systematischer Unterschied zwischen den Beobachtern wird durch Berechnung des 95%igen Konfidenzintervalls (CI) für d abgeschätzt: $d \pm t_{n-1} \cdot SE(d)$ [n = Anzahl der Studiensubjekte, SE = Standardfehler des Mittelwerts, also SE = SD/n]. Liegt Null nicht innerhalb des 95% CI, bestehen systematische Unterschiede zwischen den Beobachtern [4].

Soll bestimmt werden, ob sich in der klinischen Praxis und Forschung eine reale Veränderung des Problemausgangs ergeben hat, muss diese mindestens die Größe der kleinsten identifizierbaren Differenz (SDD) eines Messverfahrens haben. Die SDD wurde berechnet als $1,96 \times 2 \times SEM^2$ (SEM = Standardfehler der Messung, berechnet als $SD \times (1-r)$; r = Reliabilitätskoeffizient; [15]).

Da es zu den spezifischen Maßen für die Ergebnisse es keine klaren Kriterien gibt, welcher Wert eine akzeptable Interobserver-Übereinstimmung charakterisiert, wurde a priori festgelegt, dass eine Differenz der Messungen beider Beobachterinnen (d) von 10% des gesamten Messbereichs akzeptabel ist. Diese Ent-

scheidung erfolgte mehr oder weniger willkürlich, und der Leser mag bei der Interpretation unserer Ergebnisse seine eigenen Grenzen festlegen.

ICC liefern Informationen über die Fähigkeit von 2 Beobachtern, zwischen Studiensubjekten zu unterscheiden (Reliabilität; [30]). Der ICC wurde als Quotient der Varianz zwischen Patienten (Subjektvariabilität) und der Gesamtvarianz (Variabilität von Subjekt, Beobachter und zufälligem Fehler) definiert, der sich zwischen Null (keine Reliabilität) und 1 (perfekte Reliabilität) bewegt. Ein ICC von weniger als 0,75 galt als unbefriedigend [1, 18]. Die ICC wurden nach Fleiss [12] mithilfe von Zweiweg-zufallseffektmodellen berechnet.

Ergebnisse

Patienten

Die 50 Teilnehmer der Studie wurden zwischen Juli 1998 und Januar 1999 gewonnen. Ihr Durchschnittsalter (\pm SD) betrug 47 Jahre (\pm 11), 40% der Patienten waren Frauen (Tab. 1). In 35 Fällen (70%) war der dominante Arm betroffen. 9 Probanden (18%) gaben eine starke Besserung ihrer Ellenbogenbeschwerden an, aber die meisten litten immer noch unter Symptomen. 15 Patienten (30%) stufen ihre Schmerzen als mäßig bis stark ein (Schmerzen tagsüber: Einstufung \geq 3 auf einer numerischen Einstufungsskala [NRS]).

Tab. 1 Wesentliche Merkmale der Teilnehmer (IQR = Interquartilsbereich; * Modified Pain Free Function Questionnaire: Einstufungen 0–40 [40 = schwere Behinderung; 29] ** von den Patienten auf einer NRS eingestuft, Skala von 0–10 [10 = sehr schwerer Schmerz])

Merkmale	Teilnehmer (n = 50)
Weiblich (%)	20 (40)
Durchschnittsalter in Jahren (SD)	47 (11)
Dominante Seite betroffen (%)	35 (70)
Eingeschränkte Ellenbogenfunktion (Median, IQR)*	5,5 (1,5–13,3)
Schmerz tagsüber (Median, IQR)**	1 (0–4)

Übereinstimmung

Die Resultate zur Interbeobachter-Übereinstimmung hinsichtlich des Schweregrads von Beschwerden, der schmerzfreien und maximalen Griffstärke sowie der Druckschmerzschwelle am Epicondylus lateralis sind in Tab. 2 zusammengefasst. Die Einschätzungen der beiden Beobachterinnen stimmten bei allen Messungen recht gut überein. Die Schmerzintensität schien den Grad der Übereinstimmung zwischen beiden Beobachterinnen nicht zu beeinflussen (keine angegebenen Daten). Das durchschnittliche Ausmaß der Schmerzen war jedoch laut Angaben der meisten Patienten gering ($2,2 \pm 2,7$).

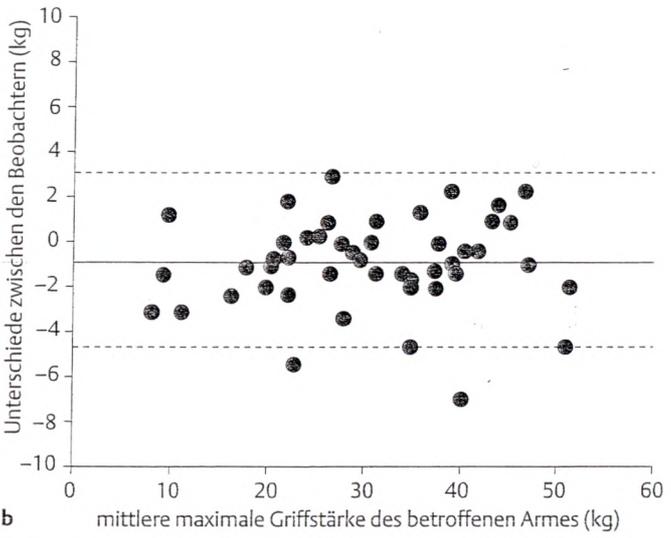
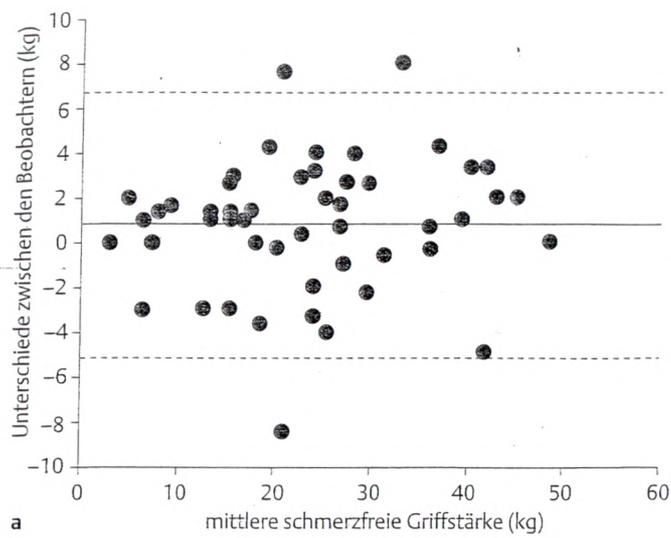
Zwischen den Beobachterinnen gab es Unterschiede im Verhältnis zum Mittelwert ihrer Beurteilungen zu schmerzfreier und maximaler Griffstärke des betroffenen Arms (Abb. 1a u. b). Die Griffstärke variierte zwischen den verschiedenen Patienten beträchtlich. Obwohl die Grenzen für Übereinstimmung bei schmerzfreier weiter als bei maximaler Griffstärke waren, fanden sich bei Letzterer systematische Unterschiede (d [95% CI] – 0,81 [–1,37 bis –0,25]). Beobachterin B gab bei etwa 72% der Patienten eine höhere maximale Griffstärke an als Beobachterin A.

Außerdem bestanden Unterschiede zwischen beiden Beobachterinnen im Verhältnis zum Mittelwert ihrer Befunde für die Druckschmerzschwelle (Abb. 2). Beobachterin B gab für ca. 74% der Patienten höhere Werte an als Beobachterin A. Mit zunehmender höherer durchschnittlicher Druckschmerzschwelle lieferte Beobachterin B häufiger höhere Werte als Beobachterin A. Die Unterschiede zwischen den Beobachterinnen nahmen auch mit zunehmenden Durchschnittswerten für den betroffenen Arm zu. Ein systematischer Fehler wurde für die Druckschmerzschwelle beim betroffenen Arm gefunden (d [95% CI] – 0,66 [–0,96 bis –0,36]).

Hinsichtlich der Griffstärke war die SDD deutlich geringer als die zuvor festgelegte akzeptable Differenz von 10% des gesamten Messbereichs. Beim Schweregrad der Beschwerden entsprach die SDD der zuvor festgelegten akzeptablen Differenz von 10% (0,7 Punkte, 10% von 7 gleich 0,7). In Bezug auf die Druckschmerzschwelle lag die SDD viel höher als die zuvor festgelegte akzeptable Differenz (1,3 Punkte, 10% von 5 gleich 0,5).

Tab. 2 Interbeobachter-Übereinstimmung und SDD für Schweregrad der Ellenbogenbeschwerden, Griffstärke und Druckschmerzschwelle (n = 50; * 95%-Grenze der Übereinstimmung [A–B] \pm 1,96 SD; † in einer NRS eingestuft: Skala 0–10 [10 = sehr starke Beschwerden] ‡ Mittelwert aus 3 Versuchen; ** Werte reichen von 1–6 kg/m² [1 = geringe Schmerzschwelle], Mittelwert aus 3 Versuchen)

Maße für die Ergebnisse	Beobachterin A		Beobachterin B		Mittelwert \pm SD Beobachterin A u. B	Mittelwert \pm SD Unterschied A–B	95%-Grenzen der Übereinstimmung*	SDD
	Mittelwert \pm SD	Bereich	Mittelwert \pm SD	Bereich				
Schweregrad†	2,02 \pm 1,7	0–7	1,98 \pm 1,6	0–6	2,0 \pm 1,6	0,04 \pm 0,8	–1,53 bis 1,97	0,7
Schmerzfreie Griffstärke (kg)‡	24 \pm 12	3–48	23 \pm 11	3–48	23,4 \pm 11,4	0,79 \pm 3,0	–5,09 bis 6,6	1,4
Maximale Griffstärke (kg)								
– betroffener Arm‡	30 \pm 11	6–50	30 \pm 11	9–53	30,1 \pm 10,8	–0,81 \pm 2,0	–4,73 bis 3,11	0,8
– nichtbetroffener Arm‡	34 \pm 10	16–54	34 \pm 10	17–52	34,3 \pm 9,8	–0,17 \pm 1,9	–3,89 bis 3,55	0,7
Druckschmerzschwelle (kg/m ²)								
– betroffener Arm**	3,3 \pm 1,5	1,2–6	3,9 \pm 1,6	1–6	3,6 \pm 1,5	–0,66 \pm 1,1	–2,82 bis 1,50	1,5
– nichtbetroffener Arm**	5,0 \pm 1,3	1,8–6	5,1 \pm 1,2	1,6–6	5,0 \pm 1,2	–0,17 \pm 0,9	–1,93 bis 1,59	1,3



Reliabilität

Die Resultate der Varianzanalyse (ANOVA) sind in Tab. 3 dargestellt. Die systematischen Unterschiede zwischen den Beobachterinnen hinsichtlich der Beurteilung von maximaler Griffstärke und Druckschmerzschwelle beim nichtbetroffenen Arm wurden durch die niedrigen p-Werte für Beobachtervarianz bestätigt. Die ICC-Werte reichten von 0,72 bis 0,98. Die Reliabilität der Beurteilungen der Druckschmerzschwelle war relativ schlecht und für den nichtbetroffenen Arm unbefriedigend (ICC < 0,75). Eine nahezu vollkommene Reliabilität ergab sich bei der schmerzfreien und der maximalen Griffstärke (ICC ≥ 0,97).

Diskussion

Die Ergebnisse zeigten hinsichtlich des Schweregrads von Beschwerden, der schmerzfreien und der maximalen Griffstärke eine ausgezeichnete Reliabilität der Beurteilungen (ICC ≥ 0,90). Bei der Druckschmerzschwelle waren die Beurteilungen deutlich weniger zuverlässig (ICC-Bereich: 0,72 – 0,77).

Systematische Unterschiede zwischen den beiden Physiotherapeutinnen ergaben sich bei der Einschätzung von maximaler Griffstärke und Druckschmerzschwelle im betroffenen Arm. Diese könnten zum Teil auf unterschiedliche verbale Ermutigung der Patienten seitens der Therapeutinnen bei der Messung der maximalen Griffstärke, Unterschiede bei den Anweisungen, Platzierung des Algometers auf dem Epicondylus lateralis, Geschwindigkeit des Druckanstiegs sowie der Reaktionszeit der Physiotherapeutinnen auf keine weitere Druckausübung zurückzuführen sein.

In den Anweisungen der *American Society of Hand Therapists* (ASHT; [8]) für eine standardisierte Positionierung des Patienten zur Messung der Griffstärke heißt es: „Der Patient sollte sitzen, die Schulter adduziert und bezüglich Rotation in Neutral-Null-Stellung, der Ellenbogen um 90° flektiert und Unterarm und Handgelenk in Neutral-Null-Stellung sein“. Möglicherweise registriert eine Messung der Griffstärke mit extendiertem Ellenbogen wie in der vorliegenden Studie diese Messgröße bei Patienten mit lateraler Epikondylitis empfindlicher [32, 33], schränkt aber die Vergleichbarkeit mit Messungen der Griffstärke bei der empfohlenen standardisierten Positionierung ein.

Weiterhin empfiehlt ASTH, dass „zur Bewertung der Griffstärke die 2. Position der Handhabung des Dynamometers benutzt [...] und der mittlere von 3 aufeinander folgenden Versuchen als Maß für den Ausgang der Griffstärkemessung verwendet werden sollte“ [8]. In der vorliegenden Studie wurde die Handhabungsposition des Jamar-Dynamometers der Handgröße des Patienten angepasst und für jede Messung standardisiert.

Frühere Studien [3, 19] hatten bei gesunden Personen und bei Patienten mit Problemen kumulativer Traumata eine ausgezeichnete Interbeobachter-Reliabilität von Griffstärkemessungen mit dem Jamar-Dynamometer ergeben (ICC-Bereich: 0,93 – 0,98). Diese Studie bestätigte die Ergebnisse für Patienten mit lateraler Epikondylitis und zeigte, dass sich die Beurteilungen von maximaler und schmerzfreier Griffstärke des betroffenen Arms durchschnittlich nur um 0,8 kg unterschieden.

Abb. 1 a u. b Unterschiede zwischen den Beobachterinnen, bezogen auf den Mittelwert der Befunde. a Schmerzfreie Griffstärke. b Maximale Griffstärke (nur beim betroffenen Arm; n = 50; — mittlere Unterschiede; - - - Grenzen der Übereinstimmung).

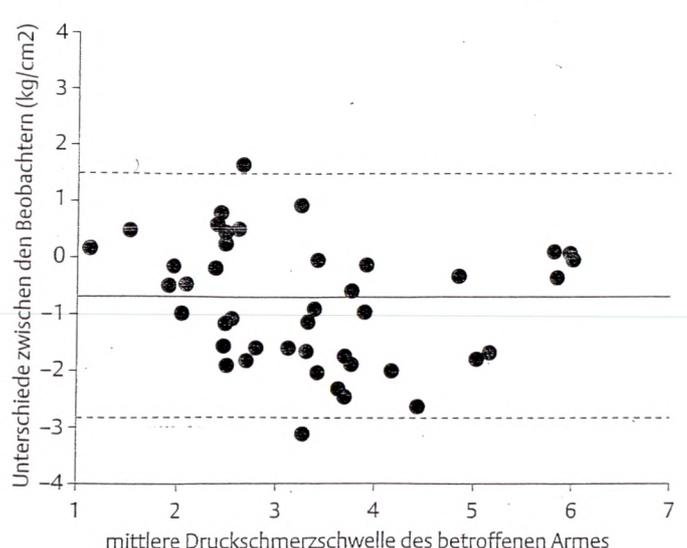


Abb. 2 Unterschiede zwischen den Beobachterinnen, bezogen auf den Mittelwert der Befunde zur Druckschmerzschwelle (nur beim betroffenen Arm; n = 50; — mittlere Unterschiede; - - - Grenzen der Übereinstimmung).

Tab. 3 Interbeobachter-Reliabilität: ANOVA und ICC für Messungen des Schweregrads der Ellenbogenbeschwerden, Griffstärke und Druckschmerzschwelle durch 2 Beobachterinnen (n = 50; NS = nicht statistisch signifikant [$p > 0,05$])

	Quelle der Variation							
	Zwischen Patienten			Zwischen Beobachterinnen			ICC	95 % CI
	Quadratisches Mittel	F	P	Quadratisches Mittel	F	P		
Schweregrad	5,31	18,62	<0,01	0,04	0,14	NS	0,90	0,83–0,94
Schmerzfremde Griffstärke	258,35	57,37	<0,01	15,73	3,49	NS	0,97	0,94–0,98
Maximale Griffstärke								
– betroffener Arm	233,60	115,34	<0,01	16,54	8,17	<0,01	0,98	0,97–0,99
– nichtbetroffener Arm	192,39	104,97	<0,01	0,69	0,38	NS	0,98	0,97–0,99
Druckschmerzschwelle								
– betroffener Arm	4,25	7,56	<0,01	10,91	19,39	<0,0001	0,77	0,62–0,86
– nichtbetroffener Arm	2,72	6,16	<0,01	0,71	1,60	NS	0,72	0,55–0,83

Um die Wirksamkeit einer Intervention zu bestimmen, müssen die Faktoren bekannt sein, die zur Variation der beobachteten Messwerte beitragen. Beim Versuch, eine reale Veränderung von statistischer Signifikanz zwischen 2 Beobachtungen zu erkennen, muss diese Veränderung mindestens so groß sein wie die SDD des Messverfahrens. Die SDD für Griffstärke bewegte sich im Rahmen der zuvor festgelegten Kriterien für Interbeobachter-Übereinstimmung (<10% des gesamten Messbereichs). Hinsichtlich der Beurteilung des Schweregrads der Probleme allgemein war die SDD genau so groß wie dieses Kriterium (0,7 Punkte entsprechen 10% von 7).

In Bezug auf die Druckschmerzschwelle lag die SDD deutlich höher als 10% (1,3 > 10% von 5), was darauf hinweist, dass sich ein echter Unterschied der Messwerte unter Umständen schwer feststellen lässt, wenn die Messungen von verschiedenen Beobachtern ausgeführt werden. Dieses Ergebnis bestätigt wiederum, dass sich die Druckschmerzschwelle als klinimetrisches Maß bei Patienten mit lateraler Epikondylitis als relativ ungünstig erweist.

Diese Studie war auf die Bewertung der Reproduzierbarkeit von Befunden zu Maßen für die Ergebnisse bei lateraler Epikondylitis beschränkt. Die Validität der Maße lässt sich schwer beurteilen und wurde selten weitergehend untersucht. Zwar gibt es wegen fehlender Goldstandards auf diesem Gebiet keine Daten zur Kriteriumsvalidität, jedoch liefern einige Studien nützliche Angaben zur Konstruktvalidität von Druckschmerzschwelle und Griffstärke bei lateraler Epikondylitis [23, 24, 29, 37]. So war die Druckschmerzschwelle bei Patienten mit lateraler Epikondylitis im Vergleich zu gesunden Personen [24] oder zum nichtbetroffenen Ellenbogen [37] signifikant niedriger. Pienimäki et al. [24] konnten jedoch keinen Zusammenhang zu Schmerzzeichnungen oder Einstufungen auf einem Fragebogen zu Schmerz und Behinderung feststellen, was die Validität der Druckschmerzschwelle als Maß bei lateraler Epikondylitis infrage zu stellen scheint.

Daten zur Konstruktvalidität von Griffstärke als Maß bei lateraler Epikondylitis sind ebenfalls rar, aber neuere Untersuchungen zeigten einen Zusammenhang zwischen schmerzfreier Griffstärke und eingeschränkter Arbeitsfähigkeit [23] sowie eingeschränkter Funktionsfähigkeit (gemessen anhand des Fragebogens zu schmerzfreier Funktionsfähigkeit und mit einer visuellen Analogskala) auf [29]. Hinsichtlich maximaler Griffstärke sind solche Zusammen-

hänge schwächer. Die Ergebnisse scheinen einleuchtend, da die Patienten versuchen, bei alltäglichen Aktivitäten Schmerz zu vermeiden und wohl selten ihre maximale Griffstärke einsetzen. Somit wird ihre funktionelle Fähigkeit hauptsächlich von ihren schmerzfreien Fähigkeiten bestimmt.

Darüber hinaus wurden Zusammenhänge zwischen Veränderungen der schmerzfreien Griffstärke im Laufe der Zeit und Veränderungen der funktionellen Einschränkung sowie der wahrgenommenen allgemeinen Besserung festgestellt. Im Gegensatz dazu fand sich kein Zusammenhang zwischen maximaler Griffstärke und einer Besserung der Symptome [32, 33]. Diese Ergebnisse und ihre klinischen Implikationen sprechen dafür, schmerzfreie Griffstärke bei lateraler Epikondylitis als Maß einzusetzen.

Die Forschungsphysiotherapeutinnen in der vorliegenden Studie wussten nichts von den Befunden der jeweils anderen Therapeutin bezüglich aller 50 Patienten. Daher gelten die Verblindung als erfolgreich und die Ergebnisse zur Interbeobachter-Reproduzierbarkeit von Befunden für unbeeinflusst vom Wissen um die Befunde der 2. Beobachterin. Die Therapeutinnen waren gut geschult, hatten umfangreiche Erfahrungen mit der Befunderhebung bei Patienten mit lateraler Epikondylitis und die Vorgehensweise zur Befunderhebung gründlich besprochen. Infolgedessen sind nicht alle Ergebnisse der Studie zur Reproduzierbarkeit von Befunden direkt auf den klinischen Alltag zu verallgemeinern.

Dagegen lässt sich das Verfahren zur Beurteilung der Griffstärke sehr genau beschreiben und ein entsprechender Befund ohne große Ausbildung erheben. Aus diesem Grund ist die externe Validität der Ergebnisse zur Griffstärke sehr zuverlässig.

Für die Beurteilung des Schweregrads trifft dies insgesamt vielleicht nicht zu. Die Studie zeigte, dass die Reproduzierbarkeit dieser Befunde zwar sehr gut, aber nicht so gut wie die der Griffstärke war, selbst bei den gut ausgebildeten Beobachterinnen. Im Rahmen üblicher medizinischer Versorgung werden die Befunde hinsichtlich dieses Maßes bei verschiedenen Untersuchern wohl noch viel stärker variieren, da sie sehr stark von der subjektiven Interpretation der Symptome und Zeichen abhängen. Diese subjektiven Interpretationen des Schweregrads von Beschwerden sind jedoch in der klinischen Praxis ein wichtiger Aspekt der Befunderhebung, sodass das Maß an sich also klinisch relevant ist. Der Leser sollte

bei der Bewertung des Schweregrads von Ellenbogenproblemen und bei der Diskussion eigener Befunde mit anderen Ärzten oder Therapeuten die potenzielle Unterschiedlichkeit der Beurteilungen verschiedener Untersucher berücksichtigen. Schmerzfreie Griffstärke eignet sich vielleicht besser, um die Schwere einer lateralen Epikondylitis zu quantifizieren, wenn für die klinische Praxis oder zu Forschungszwecken genaue Befunde gebraucht werden.

Schlussfolgerungen

Die Interbeobachter-Reproduzierbarkeit von Befunden war bei allen Maßen für die Ergebnisse mit Ausnahme der Druckschmerzschwelle ausgezeichnet. Die Gesamteinschätzung der Schwere der Symptome stimmte bei den geschulten Physiotherapeuten gut überein und war sehr verlässlich, was im klinischen Alltag ein relevantes Maß darstellt.

Eine gewisse Unterschiedlichkeit der Befunde zwischen verschiedenen Beobachtern sollte berücksichtigt werden, insbesondere wenn sich Hintergründe und Ausbildungen der Untersucher unterscheiden.

Die Beurteilungen der schmerzfreien Griffstärke zeigten sehr große Reliabilität und Übereinstimmung. Entsprechende Befunde sind relativ einfach zu erheben, und ein Zusammenhang zwischen diesem und anderen Maßen für funktionelle Einschränkung bei Patienten mit lateraler Epikondylitis ist nachgewiesen. Daher ist die Nutzung der schmerzfreien Griffstärke als Maß sowohl in der Forschung als auch in klinischer Praxis zu empfehlen.

Danksagung

Wir danken Samyra Keus für die Durchführung der Befunderhebungen.

Die Studie wurde von der *Netherlands Organization for Scientific Research* und dem *Health Insurance Council's Fund for Investigative Medicine* gefördert.

Literatur

- 1 Andresen EM. Criteria for assessing the tools of disability outcomes research. *Arch Phys Med Rehabil* 2000; 81 (Suppl 2): S15–S20
- 2 Antonaci F, Sand T, Lucas GA. Pressure algometry in healthy subjects: interexaminer variability. *Scand J Rehabil Med* 1998; 30: 3–8
- 3 Bellace JV, Healy D, Besser MP et al. Validity of the Dexter Evaluation System's Jamar dynamometer attachment for assessment of hand grip strength in a normal population. *J Hand Ther* 2000; 13: 46–51
- 4 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; 1: 307–310
- 5 Buskila D, Fefer P, Harman-Boehm I et al. Assessment of nonarticular tenderness and prevalence of fibromyalgia in hyperprolactinemic women. *J Rheumatol* 1993; 20: 2112–2115
- 6 Delaney GA, McKee AC. Inter- and intrarater reliability of the pressure pain threshold meter in measurement of myofascial trigger point sensitivity. *Am J Phys Med Rehabil* 1993; 72: 136–139
- 7 Dondt W, Willaeyts T, Verbruggen LA et al. Pain threshold in patients with rheumatoid arthritis and effect of manual oscillations. *Scand J Rheumatol* 1999; 28: 88–93
- 8 Fess EE, Moran CA, Hrsg. Clinical assessment recommendations. *American Society of Hand Therapists*. Chicago, 1981
- 9 Fess EE. A method for checking Jamar dynamometer calibration. *J Hand Ther* 1987; 1: 28–32
- 10 Fischer AA. Pressure algometry over normal muscles. Standard values, validity and reproducibility of pressure threshold. *Pain* 1987; 30: 115–126
- 11 Fischer AA. Documentation of myofascial trigger points. *Arch Phys Med Rehabil* 1988; 69: 286–291
- 12 Fleiss JF. The design and analysis of clinical experiments. London: Wiley, 1986
- 13 Geoffroy P, Yaff MJ, Rohan I. Diagnosing and treating lateral epicondylitis. *Can Fam Physician* 1994; 40: 73–78
- 14 Granges G, Littlejohn G. Pressure pain threshold in pain-free subjects, in patients with chronic regional pain syndromes, and in patients with fibromyalgia syndrome. *Arthritis Rheum* 1993; 36: 642–646
- 15 Guyatt G, Walter S, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. *Chronic Dis* 1987; 40: 171–178
- 16 Jensen K, Andersen HO, Olesen J et al. Pressure pain threshold in human temporal region. Evaluation of a new pressure algometer. *Pain* 1986; 25: 313–223
- 17 Koes BW, Bouter LM, van Mameren H et al. Randomised clinical trial of manipulative therapy and physiotherapy for persistent back and neck complaints: results of one year follow up. *BMJ* 1992; 304: 601–605
- 18 Kramer MS, Feinstein AR. Clinical biostatistics LIV. The biostatistics of concordance. *Clin Pharmacol Ther* 1981; 29: 11–23
- 19 MacDermid JC, Kramer JF, Woodbury MG et al. Interrater reliability of pinch and grip strength measurements in patients with cumulative trauma disorders. *J Hand Ther* 1994; 7: 10–14
- 20 Mathiowetz V, Weber K, Volland G et al. Reliability and validity of grip and pinch strength evaluation. *J Hand Surg (Am)* 1984; 9: 222–226
- 21 Nussbaum EL, Downes L. Reliability of clinical pressure pain algometric measurements obtained on consecutive days. *Phys Ther* 1998; 78: 160–169
- 22 Olson SL, O'Connor DP, Birmingham G et al. Tender point sensitivity, range of motion, and perceived disability in subjects with neck pain. *J Orthop Sports Phys Ther* 2000; 30: 13–20
- 23 Pienimäki T, Siira P, Vanharanta H. Muscle function of the hand, wrist and forearm in chronic lateral epicondylitis. *Eur J Phys Med Rehabil* 1997; 7: 171–178
- 24 Pienimäki T, Vanharanta H. Pain questionnaire, pain drawing, and pressure pain thresholds in chronic lateral epicondylitis. *Eur J Phys Med Rehabil* 1998; 8: 3–9
- 25 Rankin G, Stokes M. Reliability of assessment tools in rehabilitation: an illustration of appropriate statistical analyses. *Clin Rehabil* 1998; 12: 187–199
- 26 Reeves JL, Jaeger B, Graff-Redford SB. Reliability of the pressure algometer as a measure of myofascial trigger point sensitivity. *Pain* 1986; 24: 313–321
- 27 Smidt N. Conservative treatments for tennis elbow in primary care [Dissertation]. Wageningen (Netherlands): Ponsen & Looijen BV, 2001
- 28 Smidt N, van der Windt DA, Assendelft WJ et al. Corticosteroid injections, physiotherapy, or a wait-and-see policy for lateral epicondylitis: a randomised controlled trial. *Lancet* 2002; 359: 657–662
- 29 Stratford P, Levy DR, Gauldie S et al. Extensor carpi radialis tendonitis: a validation of selected outcome measurements. *Physiother Can* 1987; 39: 250–255
- 30 Stratford P. Reliability: consistency or differentiating between subjects? [Letter]. *Phys Ther* 1989; 69: 299–300
- 31 Stratford PW, Norman GR, McIntosh JM. Generalizability of grip strength measurements in patients with tennis elbow. *Phys Ther* 1989; 69: 34–39
- 32 Stratford PW, Levy DR, Gowland C. Evaluative properties of measures used to assess patients with lateral epicondylitis at the elbow. *Physiother Can* 1993; 45: 160–164
- 33 Stratford PW, Levy DR. Assessing valid change over time in patients with lateral epicondylitis at the elbow. *Clin J Sport Med* 1994; 4: 88–91
- 34 Tinsley HE, Weiss DJ. Interrater reliability and agreement of subjective judgments. *J Couns Psychol* 1975; 22: 358–376
- 35 Vicenzino B, Collins D, Wright A. The initial effects of a cervical spine manipulative physiotherapy treatment on the pain and dysfunction of lateral epicondylalgia. *Pain* 1996; 68: 69–74
- 36 van der Windt DA, Koes BW, Devillé W et al. Effectiveness of corticosteroid injections versus physiotherapy for treatment of painful stiff shoulder in primary care: randomised trial. *BMJ* 1998; 317: 1292–1296
- 37 Wright A, Thurnwald P, O'Callaghan J et al. Hyperalgesia in tennis elbow patients. *J Musculoskeletal Pain* 1994; 2: 83–97