

Übersetzung, kulturelle Adaptation und Test-Retest-Reliabilität der deutschen Version des „Headache Disability Questionnaire“

Autor:

Dimo Ivanov

Markus J.Ernst,

Wissenschaftlicher Mitarbeiter der beteiligten Institution

Korrespondenzautor: markus.ernst@zhaw.ch

Beteiligte Institutionen:

Zürcher Hochschule für Angewandte Wissenschaften,

ZHAW Institut für Physiotherapie,

Katharina-Sulzer-Platz 9,

8400 Winterthur

1 ZUSAMMENFASSUNG

2 **Einleitung:** Kopfschmerzen führen oft zu hohen Alltagseinschränkungen. Spezifische
3 Fragebögen dienen als Messinstrumente zur Erfassung solcher kopfschmerzbedingter
4 Alltagseinschränkungen. Im deutschsprachigen Raum besteht zurzeit kein ideales
5 Messinstrument, welches die Schwere der Einschränkung durch Kopfschmerzen innerhalb einer
6 akzeptablen Recall-Zeit erfasst und zuverlässig misst. Der Headache Disability Questionnaire
7 (HDQ) ist ein englischsprachiges Messinstrument, welches die Anforderungen eines
8 zuverlässigen Fragebogens für Kopfschmerzen erfüllt. Es besteht keine validierte deutsche
9 Version des HDQ. Das Ziel dieser Studie war es, den HDQ ins Deutsche zu übersetzen und
10 kulturübergreifend an die deutschsprachige Population zu adaptieren sowie seine Test-Retest-
11 Reliabilität zu untersuchen.

12 **Methodik:** In einem zwei-Phasen Querschnittstudiendesign wurde der HDQ zuerst nach
13 internationalen Richtlinien ins Deutsche übersetzt und adaptiert (HDQ-G). Zwölf erwachsene
14 deutsche Muttersprachler*innen nahmen an kognitiven Interviews zur Verständlichkeit einer
15 präfinalen deutschen Version des HDQ teil. Anschliessend wurden Anpassungen in Absprache
16 mit dem Original-Autor und einem Expertenkomitee vorgenommen. Danach wurde die Test-
17 Retest-Reliabilität des HDQ-G mittels Intraklassenkorrelationskoeffizienten (ICC) mit Hilfe von 29
18 erwachsenen Kopfschmerz-Patient*innen evaluiert und mittels Bland-Altman-Plots bildlich
19 dargestellt. Teilnehmer*innen nahmen in einem Zeitraum von zwei bis vierzehn Tagen je zweimal
20 an einer Online-Umfrage mit der finalen Version teil.

21 **Ergebnisse:** Alle Schritte des Übersetzungsprozesses wurden eingehalten und die finale
22 deutsche Version des HDQ wurde mit wenigen Ergänzungen fertiggestellt. Der HDQ-G zeigte eine
23 sehr gute Test-Retest-Reliabilität mit einem ICC=0.89 (95% Konfidenzintervall (KI) von 0.77 -
24 0.95). Die mittlere Differenz beider Umfragen lag bei +1.72 Punkten, 95% aller Werte lagen
25 zwischen +10.77 und -7.33 Punkten.

26 **Schlussfolgerung:** Die originale Version des HDQ wurde übersetzt und für die Verwendung bei
27 der erwachsenen deutschsprachigen Population im ambulanten physiotherapeutischen Setting
28 adaptiert. Der HDQ-G zeigte eine sehr gute Test-Retest Reliabilität und akzeptable
29 Übereinstimmung bei allgemeinen Kopfschmerzpatient*innen im physiotherapeutischen Setting.
30 Zusätzliche Studien sind notwendig, um weitere wissenschaftliche Gütekriterien des HDQ zu
31 untersuchen.

32

33 Die Studie wurde im Register clinicaltrials.gov registriert (clinicaltrials.gov: NCT04877561).

34 **Schlüsselwörter:** Kopfschmerzen, Assessment, Reliabilität, Übersetzung, Adaptation,
35 Headache Disability Questionnaire

36

37 **English abstract**

38 **Introduction:** Headaches are causing high disability. Specific questionnaires should serve as
39 measures for headache-related disability. In German-speaking countries, there is no ideal
40 measurement to reliably measure the severity of disability due to headache within a reasonable
41 recall time. The Headache Disability Questionnaire (HDQ) is an English measurement that fulfills
42 the requirements for a reliable headache questionnaire. It has not yet been translated into
43 German. The aim of this study was to translate the HDQ into German, to adapt it cross-culturally
44 to the German-speaking population, and to investigate its test-retest reliability.

45 **Methodology:** In a two-phase cross-sectional study design, the HDQ was first translated into
46 German and adapted according to the international guidelines (HDQ-G). Twelve adult German
47 native speakers participated in cognitive interviews on the comprehensibility of a pre-final German
48 version. Subsequently, adaptations were made and documented in agreement with the author of
49 the English version and an expert committee. Test-retest reliability of the final HDQ-G was
50 subsequently evaluated in a sample of 29 adult headache patients using intraclass-correlation-
51 coefficients (ICC) and visualized by a Bland-Altman plot. Participants completed an online survey
52 with the final version twice within a period of two to fourteen days.

53 **Results:** All steps of the translation process were followed and the final German version of the
54 HDQ was completed with few additions. The HDQ-G showed very good test-retest reliability with
55 an ICC=0.89 (95% confidence interval (CI) of 0.77 to 0.95). The mean difference of both surveys
56 was +1.72 points, with 95% of all values within +10.77 and -7.33 points.

57 **Conclusion:** The original version of the HDQ was translated and adapted for use in the adult
58 German-speaking population in the outpatient physical therapy setting. The HDQ-G showed very
59 good test-retest reliability and acceptable agreement for the general headache population in
60 physiotherapy settings. Further studies are needed to investigate other scientific criteria of the
61 HDQ.

62 The study has been registered at clinicaltrials.gov (clinicaltrials.gov: NCT04877561).

63 **Keywords:** headache, assessment, reliability, translation, adaptation, Headache Disability
64 Questionnaire

65 EINLEITUNG

66 Kopfschmerzen gehören zu den häufigsten Erkrankungen Europas, mit einer 1-Jahres Prävalenz
67 um die 20% für Migräne und gegen 30% für Spannungskopfschmerzen, zählen diese zu den
68 meistverbreiteten Kopfschmerzerkrankungen (Deuschl et al 2020). Kopfschmerzen sind in der
69 Regel sehr belastend und tragen zu starken Beeinträchtigungen der Lebensqualität von
70 Betroffenen und ihren Angehörigen bei. Im Jahre 2016 litten weltweit beinahe drei Milliarden
71 Menschen (Stovner et al. 2018), an Migräne oder Spannungskopfschmerzen (Stovner et al.
72 2018). Besonders einschränkend sind diese für Menschen im Alter von 15-49 Jahren, bei denen
73 die genannten Kopfschmerztypen zusammen die Ursache für 9.5% aller Lebensjahre mit
74 Beeinträchtigung darstellen (Stovner et al. 2018).

75 Unter den zahlreichen Möglichkeiten zur Behandlung von Kopfschmerzen haben sich auch
76 physiotherapeutische Massnahmen als nützliche Interventionen etabliert (Bendtsen & Jensen
77 2011). Zirka 25% aller Kopfschmerz-Patient*innen und 32% aller Migräne Patient*innen aus der
78 allgemeinen Population hatten zwischen 2000 und 2015 physiotherapeutische Massnahmen
79 genutzt (Moore et al. 2017). Im Sinne eines patientenzentrierten Behandlungsansatzes ist es für
80 medizinische Fachkräfte wichtig, die Auswirkungen von Kopfschmerzen auf die Lebensqualität
81 der Betroffenen zu erfassen, und die Wirkung ihrer Massnahmen mit validen Instrumenten zu
82 überprüfen. Dabei ist die Selbstdeklaration über das Vorhandensein, die Schwere, die Häufigkeit
83 und die Auswirkungen von Kopfschmerzen mitentscheidend für das Aufzeigen der Wirksamkeit
84 therapeutischer Interventionen (Haywood et al. 2018).

85 Es stehen verschiedene Fragebögen zur Verfügung, die das Ziel haben, eine patientenbasierte
86 Bewertung der Auswirkungen von Kopfschmerzen auf das Empfinden, die Körperfunktion und
87 den Alltag von Betroffenen darzulegen (Haywood et al. 2018). Allerdings weisen die bestehenden
88 Fragebögen verschiedene Defizite bei der Erfassung von Behandlungseffekten von nicht-
89 medikamentösen Interventionen bei Kopfschmerzen auf (Luedtke et al. 2020).

90 Aktuell werden in klinischen Settings besonders zwei Kopfschmerzfragebögen empfohlen
91 (Haywood et al. 2018; Luedtke et al. 2020), der „Migräne disability assessment“ (MIDAS)
92 Fragebogen ist bei der Diagnostik und Behandlung von Migräne in vielen Ländern bereits etabliert
93 (Benz et al. 2018). Trotz guter Praktikabilität und psychometrischen Eigenschaften wurden
94 allerdings auch einzelne Aspekte des Fragebogens kritisiert. Zum einen wird seine
95 Durchführbarkeit aufgrund der Aufforderung zu einer exakten Angabe von Tagen mit
96 Beeinträchtigungen durch Kopfschmerzen als zu anspruchsvoll für Patient*innen erachtet.

97 Darüber hinaus könnte das Abrufen eines 3-monatigen Zeitraums zu Verzerrungen im Sinne
98 eines eingeschränkten «Recalls» führen und somit die Präzision der Messung verringern (Benz
99 et al. 2018; Luedtke et al. 2020).

100 Der „Headache Impact Test“ (HIT-6) dagegen fordert die Befragten dazu auf, sich nur an einen
101 Zeitraum von einem Monat zu erinnern, und scheint daher besser geeignet zu sein als der MIDAS
102 (Kosinski et al. 2003; Luedtke et al. 2020). Dieser Fragebogen wurde bereits in neun Sprachen
103 validiert und steht auch im deutschsprachigen Raum zur Verfügung (Martin et al. 2004). Obwohl
104 der HIT-6 beim Screening von Kopfschmerzen empfohlen wird, zeigt er Schwächen bei der
105 Erfassung und Berechnung des Ausmasses einer Beeinträchtigung (Score zwischen 36-78
106 Punkten) und deren Auswirkung auf die Lebensqualität von Betroffenen (Luedtke et al. 2020;
107 Niere 2009). Das Bewertungssystem bietet wenig Auswahlmöglichkeiten und die
108 Punkteverteilung kann zu Schwierigkeiten bei der Interpretation der Ergebnisse führen (Niere
109 2009). Daher besteht weiterhin Bedarf nach einem zuverlässigen Fragebogen, der ausführlich
110 die Schwere der Beeinträchtigungen durch Kopfschmerzen misst und eine akzeptable «Recall-
111 Zeit» erfasst (Luedtke et al. 2020).

112 Der „Headache disability questionnaire“ (HDQ) ist ein bisher noch nicht validierter Fragebogen
113 für den deutschsprachigen Raum. Er stellt ein Messinstrument dar, welches den
114 Behandlungseffekt von nicht-medikamentösen Interventionen bei Kopfschmerzen messen kann,
115 und ähnlich wie der HIT-6 eine vierwöchige Recall-Zeit hat, jedoch ein einfacheres
116 Auswertungssystem bietet (Niere 2009). Der Fragebogen wurde für die Erhebung von Daten zur
117 Lebensqualität der allgemeinen Kopfschmerzpopulation in physiotherapeutischer Behandlung
118 entwickelt (Niere & Quin 2009). Er besteht aus neun intervallskalierten Items, die mittels einer
119 Skala von 0 bis 10 beurteilt werden können und einen Gesamtscore zwischen 0-90 Punkten
120 ergeben. Dadurch wird auch eine prozentuale Angabe der Beeinträchtigung als Gesamtergebnis
121 ermöglicht. Der HDQ hebt sich von den erwähnten Fragebögen ab, indem er neben
122 Schmerzintensität auch explizit den Schweregrad der Einschränkung von alltäglichen Aktivitäten
123 im letzten Monat aufgrund von Kopfschmerzen erfasst (Niere & Quin 2009). Eine Übersetzung
124 und Anpassung für dessen Nutzung im deutschsprachigen Raum bestehen aktuell nicht.

125 Das Ziel dieser Studie war es, den HDQ-Fragebogen ins Deutsche zu übersetzen und an die
126 deutschsprachige Population zu adaptieren. Anschliessend wurde dessen Test-Retest-
127 Reliabilität untersucht.

128 METHODE

129 Die Studie verwendet ein zwei-Phasen deskriptives Querschnittsdesign.

130 Zuerst wurde der HDQ nach empfohlenen internationalen Richtlinien (Mohler et al. 2016) vom
131 Englischen ins Deutsche übersetzt und adaptiert.

132 In einem zweiten Schritt wurde die Test-Retest-Reliabilität anhand des Interklassenkorrelations-
133 koeffizienten (ICC) und der „Limits of Agreement (LoA)“ quantifiziert.

134 ÜBERSETZUNG UND KULTURELLE ADAPTATION

135 Die Übersetzung und kulturelle Adaptation wurde nach dem internationalen TRAPD Team
136 Translation Model durchgeführt (Mohler et al. 2016) (**siehe Abb. 1**).

Abb. 1: TRAPD Team Translation Model (Mohler et al., 2016)



137

138

139 **Übersetzungsprozess (Translation)**

140 Im ersten Schritt des TRAPD-Modells (Mohler et al. 2016) fand die Übersetzung des originalen
141 Fragebogens durch eine professionelle Übersetzerin statt. Diese hatte ausführliche Informationen
142 zum Gebrauch des ursprünglichen und des neuen Fragebogens erhalten. Anschliessend wurde
143 der übersetzte Fragebogen von einem Fachexperten mit deutscher Muttersprache begutachtet
144 und angepasst. In einem zweiten Schritt überprüfte eine professionelle Revisorin die Version des
145 Fachexperten aus linguistischer Sicht, woraus eine dritte Version mit entsprechenden
146 Korrekturen entstand.

147 **Expertenkomitee-Beurteilung (Review/Adjudikation)**

148 Alle drei Versionen wurden in einem Komitee aus Expert*innen mit allen bisher Beteiligten
149 diskutiert. Zusätzlich wurde der Entwickler des Original-Fragebogens kontaktiert, um
150 Verständnisfragen zu klären und die diskutierten Anpassungsvorschläge zu kommunizieren. Das
151 Expert*innenkomitee überprüfte und diskutierte sorgfältig alle Übersetzungen und
152 Unstimmigkeiten, bis ein Konsens über eine vorläufige Version des Fragebogens erzielt wurde.

153

154 **Prätest–Phase/Dokumentation**

155 In der Prätest-Phase wurden separate kognitive Interviews zur Verständlichkeit der übersetzten
156 Version durchgeführt (Pohontsch & Meyer 2015). Dabei wurden Teilnehmer*innen ausführlich zu
157 Ihrem Verständnis von Einzelfragen, aber auch des Gesamtfragebogens interviewt. Dafür wurden
158 Teilnehmer*innen mit deutscher Muttersprache, die mindestens 18 Jahre alt sind, an
159 Kopfschmerzen leiden und in der Schweiz wohnhaft sind aus 2-3 Physiotherapie-Praxen im
160 Gross-Raum Zürich/Winterthur eingeschlossen. Die Stichprobengrösse für die kognitiven
161 Interviews durfte je nach Sättigungsgrad zwischen fünf und 15 Teilnehmer*innen betragen (Willis
162 2005).

163 Personen, welche den Ein- und Ausschlusskriterien entsprachen, erhielten einen Online-Zugang
164 zu Studieninformationen, inklusive einer Einverständniserklärung sowie einer Umfrage mit der
165 übersetzten präfinalen Fragebogen-Version. Eine Sättigung wurde als erreicht betrachtet, wenn
166 keine neuen Aussagen zum Gesamt- Fragebogen oder einzelner Items seitens der
167 Teilnehmer*innen gemacht wurden.

168 Anschliessend wurden kognitive Interviews mit den Teilnehmer*innen durchgeführt. Diese
169 beinhalteten deren Einschätzungen zum Gesamteindruck des Fragebogens, zur Verständlichkeit
170 einzelner Fragen und möglichen Verbesserungsvorschlägen. Anhand der gesammelten
171 Informationen wurden durch Konsens im Expert*innenkomitee und nach Absprache mit dem
172 Original-autor nötige Anpassungen getätigt und eine finale Version (HDQ-G) für die Verwendung
173 im klinischen Setting fertiggestellt (**siehe Abb. 2**). Das gesamte Prozedere sowie die Entstehung
174 der finalen Übersetzung wurden dokumentiert.

175

Abb. 2: Finale HDQ-G Version

HEADACHE DISABILITY QUESTIONNAIRE – DEUTSCHE VERSION

Name: Datum:/...../..... Punkte / 90

Bitte lesen Sie jede Frage aufmerksam durch und kreisen Sie die Antwort ein, die am besten auf Sie zutrifft.

1. Wie würden Sie die übliche Stärke Ihrer Kopfschmerzen auf einer Skala von 0 bis 10 einschätzen?

0	1	2	3	4	5	6	7	8	9	10	SCHLIMMSTE SCHMERZEN
KEINE SCHMERZEN											

2. Wenn Sie Kopfschmerzen haben, in wie viel Prozent der Fälle, sind diese Schmerzen stark?

NIE	1-10%	11-20%	21-30%	31-40%	41-50%	51-60%	61-70%	71-80%	81-90%	91-100%	IMMER
0	1	2	3	4	5	6	7	8	9	10	

3. An wie vielen Tagen im letzten Monat haben Sie sich für eine Stunde oder länger wegen Ihrer Kopfschmerzen hingelegt?

NIE	1-3	4-6	7-9	10-12	13-15	16-18	19-21	22-24	25-27	28-31	JEDEN TAG
0	1	2	3	4	5	6	7	8	9	10	

4. Wenn Sie Kopfschmerzen haben, wie oft müssen Sie deswegen der Arbeit, dem Studium oder der Schule den ganzen Tag oder einen Teil des Tages fernbleiben?

NIE	1-10%	11-20%	21-30%	31-40%	41-50%	51-60%	61-70%	71-80%	81-90%	91-100%	IMMER
0	1	2	3	4	5	6	7	8	9	10	

5. Wenn Sie während der Arbeit (bzw. während Studium oder Schule) Kopfschmerzen haben, wie stark ist dann Ihre Arbeitsfähigkeit eingeschränkt?

NICHT	1-10%	11-20%	21-30%	31-40%	41-50%	51-60%	61-70%	71-80%	81-90%	91-100%	ARBEITS- UNFÄHIG
0	1	2	3	4	5	6	7	8	9	10	
EINGESCHRÄNKT											

6. An wie vielen Tagen im letzten Monat wurden Sie mindestens den halben Tag wegen Ihrer Kopfschmerzen davon abgehalten, haushaltsbezogene Arbeiten (inkl. Einkauf, Reparaturen, Gartenarbeit etc.) durchzuführen?

NIE	1-3	4-6	7-9	10-12	13-15	16-18	19-21	22-24	25-27	28-31	JEDEN TAG
0	1	2	3	4	5	6	7	8	9	10	

7. Wenn Sie Kopfschmerzen haben, wie stark ist dann Ihre Fähigkeit eingeschränkt, haushaltsbezogene Arbeiten (inkl. Einkauf, Reparaturen, Gartenarbeit etc.) durchzuführen?

NICHT	1-10%	11-20%	21-30%	31-40%	41-50%	51-60%	61-70%	71-80%	81-90%	91-100%	UNFÄHIG
0	1	2	3	4	5	6	7	8	9	10	
EINGESCHRÄNKT											

8. An wie vielen Tagen im letzten Monat wurden Sie durch Ihre Kopfschmerzen von nicht arbeitsbezogenen Aktivitäten (Familie, soziale Kontakte, Freizeit) abgehalten?

NIE	1-3	4-6	7-9	10-12	13-15	16-18	19-21	22-24	25-27	28-31	JEDEN TAG
0	1	2	3	4	5	6	7	8	9	10	

9. Wenn Sie Kopfschmerzen haben, wie stark ist dann Ihre Fähigkeit eingeschränkt, nicht arbeitsbezogene Aktivitäten (Familie, soziale Kontakte, Freizeit) auszuüben?

NICHT	1-10%	11-20%	21-30%	31-40%	41-50%	51-60%	61-70%	71-80%	81-90%	91-100%	UNFÄHIG
0	1	2	3	4	5	6	7	8	9	10	
EINGESCHRÄNKT											

177 TEST-RETEST-RELIABILITÄT

178 Die Test-Retest-Reliabilität des HDQ-G wurde untersucht, um die Reproduzierbarkeit der
179 Testergebnisse bei einem möglichst unveränderten Zustand von Kopfschmerzpatient*innen über
180 einen kurzen Zeitraum hinweg zu beurteilen. Dabei wurden die Ergebnisse zweier Umfragen
181 durch die gleichen Teilnehmer*innen miteinander verglichen. Dies entspricht den international
182 anerkannten Standards zu Reliabilität und Messfehler nach COSMIN (Mokkink et al., 2019).

183 **Eignungskriterien der Teilnehmer*innen**

184 Erwachsene Teilnehmer*innen, welche wegen Kopfschmerzen seit maximal drei Monate in
185 physiotherapeutischer Behandlung waren, wurden in ambulanten physiotherapeutischen
186 Privatpraxen und mit Hilfe des Schweizerischen Verbandes für orthopädisch muskuloskelettale
187 Physiotherapie (SVOMP, <https://www.svomp.ch>) in der deutschsprachigen Schweiz rekrutiert.
188 Die Kopfschmerzen mussten seit mindestens einem Monat bestehen und mindestens einmal im
189 Monat auftreten. Die typischerweise abfallende Prävalenz von Kopfschmerzen ab dem Alter von
190 65 Jahren wurde bei den Einschlusskriterien berücksichtigt (Stovner et al. 2018), weswegen nur
191 Teilnehmer*innen im Alter zwischen 18 und 65 Jahren eingeschlossen wurden. Voraussetzungen
192 zur Teilnahme an der Studie waren ein Internetzugang sowie fließende schriftliche und
193 mündliche Verständigung in deutscher Sprache. Die Eignungskriterien und die
194 Einverständniserklärung wurden zu Beginn der Online-Umfrage, direkt nach den
195 Studieninformationen, abgefragt und eingeholt. Das „Nicht-erfüllen“ eines Eignungskriteriums
196 führte automatisch zum Ausschluss der Person und dem Ende der Umfrage.

197

198 **Prozedere**

199 Die Teilnahme an der Studie wurde mit dem Online-Tool RedCap durchgeführt
200 (<https://redcap.zhaw.ch/>). Informationsflyer zur Studie mit den aufgelisteten Eignungskriterien
201 und einem QR-Code oder Web-Link zur Online-Umfrage wurden in ambulanten
202 Physiotherapiepraxen ausgelegt und an Kopfschmerzpatient*innen ausgeteilt. Mit dem QR-Code/
203 Web-Link wurde potenziell geeigneten Teilnehmer*innen Zugriff zu den Studieninformationen
204 ermöglicht. Bei Erfüllung aller Einschlusskriterien inklusive Einverständniserklärung durften sie
205 an der Online-Umfrage teilnehmen.

206 Zur Beschreibung der wahrscheinlichsten Kopfschmerzerkrankung jeder/s Teilnehmenden,
207 wurde mit einer initialen Filteruntersuchung, angelehnt an die ICHD (Internationale Klassifikation
208 von Kopfschmerzerkrankungen, 3. Auflage) und mit Fragen zur Lokalisation, Schmerzqualität,
209 Kopfschmerzfrequenz, Kopfschmerzdauer und Kopfschmerzintensität, sowie typischen

210 Begleitsymptomen wie z.B. Übelkeit oder Lichtempfindlichkeit, für die häufigsten Kopfschmerz-
211 Typen (im Besonderen Migräne, Spannungskopfschmerz) gescreent (International Headache
212 Society 2018). Anschließend füllten die Teilnehmer*innen die Fragen des HDQ-G aus. Nach 48
213 – 72 Stunden wurden die Teilnehmer*innen per E-Mail zum nochmaligen Ausfüllen des HDQ-G
214 eingeladen. Zusätzlich wurden sie aufgefordert, eine mögliche Veränderung ihrer
215 Kopfschmerzsituation gegenüber der ersten Umfrage auf einer Likert-Skala mit fünf
216 Antwortmöglichkeiten von -2, „viel schlechter“, bis +2, „viel besser“, anzugeben.
217 Teilnehmer*innen, die eine Veränderung ihres Zustandes bzw. ihrer Kopfschmerzen angegeben
218 hatten, wurden von der Analyse ausgeschlossen. Eingeschlossen wurden nur Proband*innen,
219 welche die Frage mit der Option „unverändert“ beantwortet hatten. Es wurden keine
220 Einschränkungen bzgl. der Einnahme regulärer Medikamente und Therapien gemacht.
221 Teilnehmer*innen blieben verblindet gegenüber ihren Antworten in der ersten Umfrage.
222 Ausserdem wurden die Fragen des HDQ-G bei RedCap mit Dropdownlisten gestaltet, sodass
223 den Teilnehmer*innen auch kein visueller Überblick über ihre Antworten bei den vorherigen
224 Fragen ermöglicht wurde.

225 **Datenanalyse**

226 Zur Darstellung demographischer Daten wurden deskriptive Statistiken mit Mittelwerten,
227 Standardabweichungen und Häufigkeiten verwendet. Für die Test-Retest Reliabilität wurde ein
228 Intraklassenkorrelationskoeffizient (ICC 2,1) mit einem 95% Konfidenzintervall (KI) berechnet
229 (Koo & Li, 2016). Der ICC-Wert variiert zwischen 0 und 1, je höher der Wert, desto höher die
230 Zuverlässigkeit. Explizit wurde der ICC-Wert wie folgt klassifiziert: "sehr gut" (≥ 0.81), "gut" (0.61
231 - 0.80), "mässig" (0.41 - 0.60), "schlecht" (≤ 0.40) (Landis & Koch 1977).

232 Die mittleren Differenzen zwischen beiden Erhebungen wurden mit der Bland-Altman-Analyse
233 berechnet und grafisch dargestellt. Neben der mittleren Differenz werden die Grenzen der
234 Übereinstimmung mit der knapp doppelten (1.96) Standardabweichung als obere und untere
235 Grenzen dargestellt, wobei 95% aller Differenzen innerhalb dieser Grenzen liegen (Bland &
236 Altman, 1986). Die statistische Analyse wurde mittels Software R, Version 3.6.3, inklusive der
237 «packages» „psych“ und „BlandAltmanLeh“ durchgeführt.

238

239 **Fallzahl-Berechnung**

240 Eine genügende Test-Retest-Reliabilität wurde für einen ICC ≥ 0.90 mit unterem KI von 0.7 (gilt
241 als mindestens noch «gut») festgelegt. Für eine Power von 90%, einem Typ I Fehler $\alpha = 0.05$,

242 wurde eine Mindeststichprobengröße von 25 Patient*innen a priori berechnet (Bujang, 2017).
243 Ausgehend von einer 20%-igen Drop-out-Rate nach der ersten Messung, wurde die erforderliche
244 Stichprobengröße auf $n = 32$ erhöht (Bujang 2017).

245

246 ERGEBNISSE

247 ÜBERSETZUNG UND KULTURELLE ADAPTATION

248 Die Übersetzungs- und Adaptationsprozesse wurden wie geplant durchgeführt. Die präfinale
249 Version wurde durch Einigung im Expert*innenkomitee bestimmt, wobei Verständnisunterschiede
250 nach der Übersetzung diskutiert und Anpassungen vorgenommen wurden.

251 12 Proband*innen nahmen an kognitiven Interviews teil, bei denen sie über das Verständnis und
252 den Aufwand des gesamten Fragebogens und einzelner items befragt wurden. Die
253 Teilnehmer*innen erwähnten ein gutes Verständnis zu den meisten Fragen. Sie berichteten von
254 einem geringen zeitlichen Aufwand (ca. 10 Minuten) zum Ausfüllen des Fragebogens.

255 Sieben Teilnehmer*innen gaben an, Schwierigkeiten beim Verstehen der zweiten Frage
256 (Kopfschmerz-Stärke) zu haben. Konkret nannten sie die Frageformulierung ungewöhnlich und
257 empfanden, dass diese Frage nach Häufigkeiten eines Phänomens fragt und deshalb eher nicht
258 mit Prozentangaben zu beantworten sei.

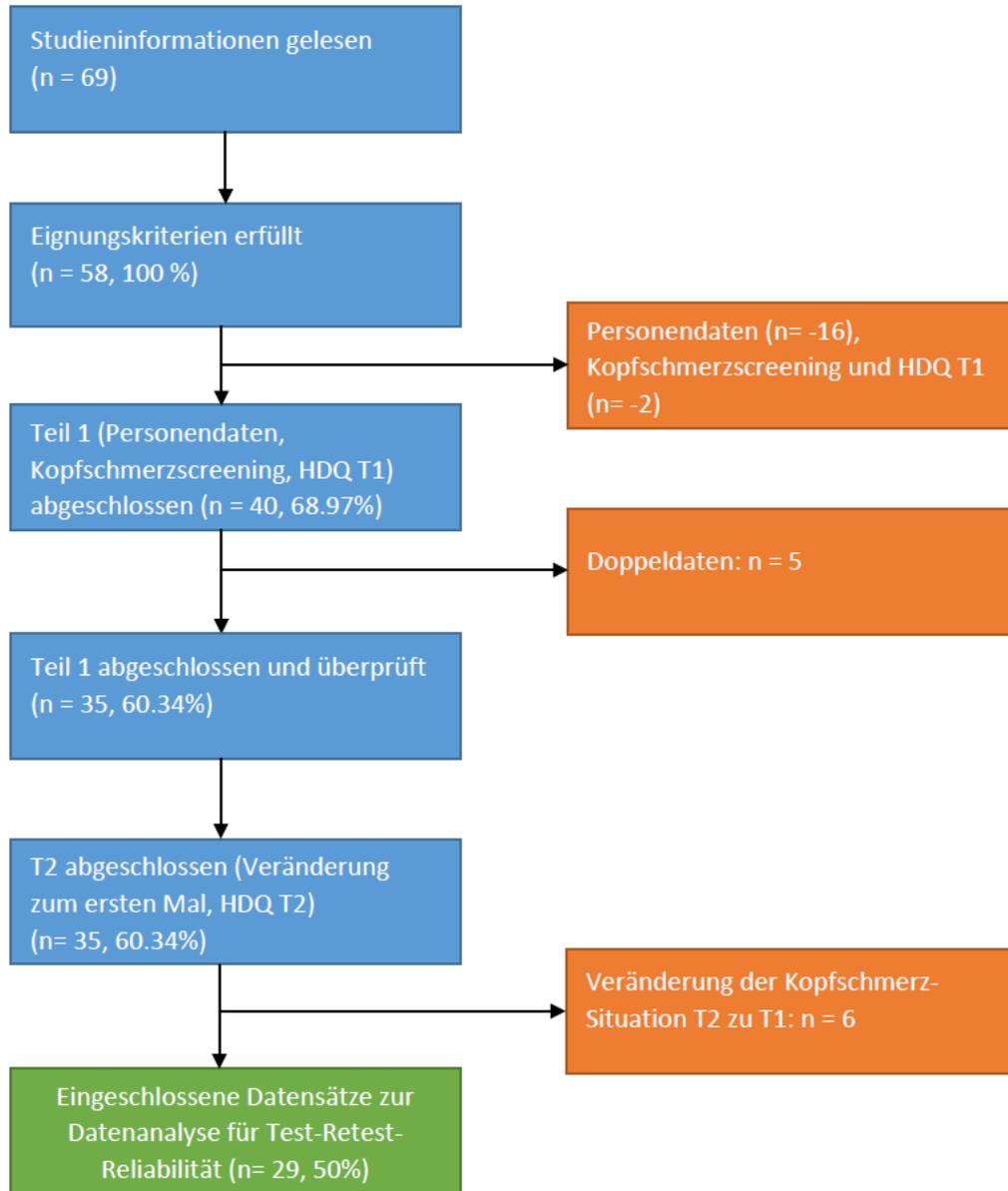
259 Ähnliche Verständnisprobleme hatten vier Teilnehmer*innen bei Frage vier (Auswirkungen von
260 Kopfschmerzen auf Arbeit, Studium, Schule), wobei für sie die Antwortmöglichkeiten in Prozenten
261 ebenfalls nicht intuitiv waren.

262 Zwei Teilnehmer*innen verstanden bei den ersten zwei Fragen nicht, auf was für einen Zeitraum
263 sich diese beziehen.

264 Weiter fanden es drei Proband*innen wichtig, dass auch die Medikamenteneinnahme für den
265 Fragebogen berücksichtigt werden sollte.

266 Die zweite Frage der präfinalen Version des HDQ wurde nach Absprache im Expert*innenkomitee
267 mit dem Zusatz „(...), „in wie viel Prozent der Fälle“, (...)“ ergänzt. Anschließend wurde die finale
268 Version des HDQ-G festgelegt (**Abb. 2**) und für die Test-Retest Reliabilität im Online-Tool
269 angepasst.

Abb. 3: Flussdiagramm des Einschlussverfahrens



270

271 TEST-RETEST-RELIABILITÄT

272 **Stichprobe**

273 Die Stichprobengröße sowie die Drop-out-Quote wurden in einem Flussdiagramm (**siehe Abb.**
274 **3**) dargestellt. Gesamt erfüllten 58 Personen die Eignungskriterien und erklärten sich zur
275 Teilnahme an der Studie bereit. Davon gaben 42 Teilnehmer*innen ihre Personendaten (Name,
276 Geschlecht, Geburtsdatum etc.) an und 40 füllten das Kopfschmerzscreening und den HDQ-G

277 erstmalig und vollständig aus. Fünf doppelte Datensätze (Personen hatten irrtümlich zweimal die
278 Umfrage ausgefüllt) wurden entdeckt, wobei jeweils nur die aktuelleren Datensätze
279 eingeschlossen blieben. Alle 35 Teilnehmer*innen füllten den HDQ-G Fragebögen innerhalb von
280 max. zwei Wochen ein zweites Mal aus (Mittelwert $4.55 \pm$ Standardabweichung (SA) 4.33 Tage),
281 wobei sechs Teilnehmer*innen für die Analyse zur Test-Retest Reliabilität der Ergebnisse
282 ausgeschlossen wurden, da sie eine Veränderung ihrer Kopfschmerzsituation (auf der Skala von
283 -2, viel schlechter, bis +2, viel besser) zwischen dem ersten und dem zweiten Ausfüllen des HDQ-
284 G, angaben. Es bestanden keine signifikanten Unterschiede zwischen den kompletten (n= 35)
285 und inkompletten (n=29) Datensätzen (**Tabelle 1**).

286

287

288 **Tabelle 1:** Demografische Daten der Test-Retest-Reliabilität Untersuchung

		T1 - MESSUNG	T2 - FOLLOW UP	P-WERTE
TEILNEHMER*INNEN ANZAHL		35	29	
Alter		38.34 ± 11.15	37.45 ± 11.66	0.76
GESCHLECHT	männlich	5 (14.29%)	4 (13.79%)	1
	weiblich	30 (85.71%)	25 (86.21%)	
KOPFSCHMERZDAUER	In Jahren	6.91 ± 9.60	7.29 ± 9.51	0.87
TEILNEHMER*INNEN MIT BEKANNTER ÄRZTLICHER DIAGNOSE	Migräne	10 (28.57%)	10 (34.48%)	0.21
	Spannungs-Kopfschmerz	2 (5.71%)	2 (6.90%)	
	Andere	5 (14.29%)	3 (10.34%)	
	Keine Angaben	18 (51.43%)	14 (48.28%)	
„WAHRSCHEINLICHER“ KOPFSCHMERZ-TYP NACH ICHD SCREENING	Migräne	19 (54.29%)	17 (58.62%)	0.20
	Spannungs-Kopfschmerz	10 (28.57%)	8 (27.59%)	
	Andere	6 (17.14%)	3 (10.34%)	
HDQ-G GESAMTERGEBNIS		28.40 ± 9.64	26.21 ± 10.40	0.39
EINNAHME VON MEDIKAMENTEN		31 (88.57%)	26 (89.66%)	
MEDIKATION HÄUFIGKEIT	Täglich	6 (17.14%)	6 (20.69%)	0.24
	Wöchentlich	8 (22.86%)	7 (24.14%)	
	Monatlich	6 (17.14%)	5 (17.24%)	
	Bei Bedarf	11 (31.43%)	8 (27.59%)	
ART DER MEDIKAMENTE	NSARs	22 (62.86%)	21 (72.41%)	1
	Spezifische Medikamente (z.B. Triptane, Lyrika, Zomik)	9 (25.71%)	8 (27.59%)	

289 *Legende Tabelle 1: HDQ-G= deutsche Version des Headache Disability Questionnaire, ICHD= Internationale*
 290 *Klassifizierung von Kopfschmerzerkrankungen, NSAR= Nicht steroidale Antiphlogistika; Werte sind Mittelwerte ±*
 291 *Standardabweichung, oder % Angaben.*

292

293 Zur Beurteilung der Reproduzierbarkeit wurden die Datensätze der 29 Teilnehmer*innen
 294 analysiert, deren Zustand innerhalb beider Messzeitpunkte als stabil bezeichnet wurde. Das

295 durchschnittliche Intervall zwischen dem Ausfüllen beider Umfragen betrug 4.37
296 (Standardabweichung (SA)=3.30) Tage.

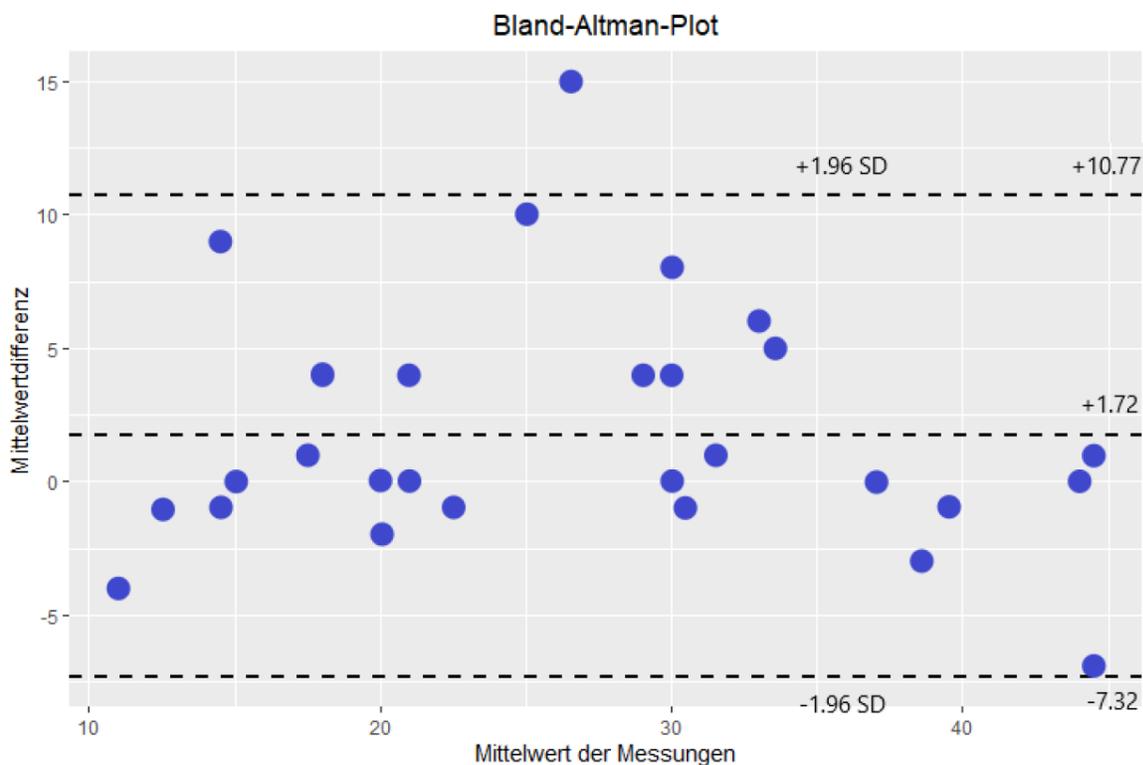
297

298 **Reliabilität**

299 Eine sehr gute Test-Retest-Reliabilität wurde für den HDQ-G mit einem ICC von 0.89 und einem
300 95% Konfidenzintervall (KI) erreicht, das sich von 0.77 bis 0.95 erstreckt.

301 Der Bland-Altman-Plot (siehe **Abb. 4**) zeigte, dass die mittlere Differenz der-Gesamtpunktzahl
302 zwischen beiden Messungen bei einem Wert von +1.72 Punkten lag. Die Grenzen der
303 Übereinstimmung (LoA) in denen sich 95% aller Differenzen befinden, lagen bei +10.77 und -7.33
304 Punkten.

Abb. 4: Bland-Altman-Plot des HDQ-G



305

306

307

308 DISKUSSION

309 Diese Studie besteht aus zwei Phasen, wobei zunächst der HDQ übersetzt und die deutsche
310 Version (HDQ-G) kulturell adaptiert wurden. Im zweiten Schritt wurde die Test-Retest Reliabilität
311 und Übereinstimmung eines zweimaligen Ausfüllens der finalen Version des HDQ-G-
312 Fragebogens evaluiert (Giavarina 2015; McGraw 2017).

313 **Übersetzungsprozess des HDQ zum HDQ-G**

314 Das Vorgehen bei der Übersetzung und Adaptation nach dem TRAPD-Modell zeigte eine neue
315 Perspektive des Übersetzungsprozesses und verspricht eine professionelle Herangehensweise
316 mit geringem Aufwand (Mohler et al. 2016). Gemäss dem TRAPD-Modell ist eine
317 Rückübersetzung überflüssig und sogar ungünstig für das Endergebnis, den deutschsprachigen
318 Fragebogen, denn dadurch kann die eigentliche Übersetzung an Natürlichkeit verlieren und als
319 schwerer verständlich oder irrtümlich als falsch bezeichnet werden (Behr 2017). Epstein et al.
320 (2015) verglichen verschiedene Übersetzungsansätze und lieferten gute Evidenz, dass eine
321 Rückübersetzung nur geringe Veränderungen erzeugt, und unterlassen werden kann. Viel
322 wichtiger sei ein Expert*innenkomitee mit guten Kultur- und Sprachkenntnissen beider Sprachen,
323 was zur Sicherstellung des korrekten Inhalts beiträgt (Epstein et al. 2015). Beim TRAPD-Modell
324 wird die Sicherstellung der Qualität der Übersetzung durch die Einbeziehung geeigneter
325 Übersetzungs-, Inhalts- und Umfrageexpert*innen in einem Teamansatz ermöglicht und eine
326 gründliche Dokumentation durchgeführt (Mohler et al. 2016).

327 Der Übersetzungsprozess wurde bis zur Prätest-Phase planmässig durchgeführt und die Review-
328 Beteiligten waren sich betreffend einer präfinalen Version des HDQ-G einig. Jedoch wurden
329 einige potenzielle Verständnisprobleme während den kognitiven Interviews seitens der beteiligten
330 Kopfschmerzpatient*innen angesprochen. Für einen grösseren Anteil der Teilnehmer*innen
331 waren *%-Angaben* zur Häufigkeit von starken Kopfschmerzen (Frage 2) und Auswirkungen von
332 Kopfschmerzen auf Arbeit, Studium, Schule (Frage 4) schwer verständlich. Sie äusserten, dass
333 sie Antwortmöglichkeiten „in Tagen“ bevorzugten. Beide Fragen beginnen mit dem gleichen
334 Nebensatz und haben den gleichen Anfang des Hauptsatzes: „*Wenn sie Kopfschmerzen haben,*
335 *wie oft (...)*“, übersetzt aus dem Englischen: „*When you have headache, how often (...)*“. Eine
336 Vermutung, warum die Teilnehmerinnen Schwierigkeiten hatten, die zwei Fragen in Prozent zu
337 beantworten, war, dass der klassische Satzbau missachtet wurde. Bei Konditionalsätzen wird im
338 Nebensatz die Bedingung beschrieben und im Hauptsatz die Folge. Wird dies nicht erkannt, kann

339 der Satz eine andere Bedeutung annehmen (Klein 1993). Daher wurde die zweite Frage nach
340 Absprache im Expert*innenkomitee angepasst: „..., in wie viel Prozent der Fälle...“.

341 Ein weiterer Einwand aus den Interviews betraf die ersten zwei Fragen. So war es für die
342 Teilnehmer*innen unverständlich, auf welche Zeitperiode sich beide Fragen beziehen. Nach
343 Rücksprache mit dem Autor der Originalversion wurde jedoch im Expert*innenkomitee
344 entschieden, keine Veränderungen diesbezüglich vorzunehmen, da eine Ergänzung mit zeitlicher
345 Einrahmung eine zu große Veränderung des Original-Fragebogens HDQ bedeutet. Bei der
346 dritten, sechsten und achten Frage des HDQ-G wird nach der Häufigkeit in Tagen innerhalb eines
347 Zeitraums von einem Monat gefragt. Die Autoren der Originalversion strebten eine Heterogenität
348 der Fragen für häufigkeitsbezogene und nicht-häufigkeitsbezogene Elemente an (Nieme & Quin
349 2009). Ähnliches gilt auch für den HIT-6, wo Zeiträume erst ab der vierten Frage erwähnt werden
350 (Kosinski et al. 2003). Beim MIDAS dagegen wird der Zeitraum für den „Recall“ vorgegeben
351 (Stewart et al. 2001). Für die Zukunft bleibt es offen, ob inhaltliche Veränderungen in Bezug auf
352 den Recall-Zeitraum am HDQ-G getätigt werden sollten.

353 Die Erfassung der Einnahme von Kopfschmerzmedikamenten wurde ebenfalls als eine
354 potenzielle inhaltliche Ergänzung zum HDQ-G-Fragebogen von drei Teilnehmer*innen erörtert,
355 weil eine Medikamenteneinnahme auch einen Einfluss auf die Dauer und Intensität ihrer
356 Kopfschmerzen habe. Die Einnahme von Kopfschmerz-Medikamenten wird auch in vielen nicht-
357 medikamentösen Kopfschmerzstudien als Outcome-Messung erfasst (Amons et al. 2019;
358 Andersen et al. 2017; Muñoz-Gómez et al. 2021). Auch in dieser Studie gab eine grosse Anzahl
359 Teilnehmer*innen an regelmässig Medikamente gegen Kopfschmerz einzunehmen (Tabelle1).
360 Dennoch scheint die Einbeziehung von medikamentenspezifischen Elementen in
361 Kopfschmerzfragebögen nicht üblich zu sein (Luedtke et al. 2020). Gemäß Originalautor des HDQ
362 sind die Informationen zum Medikamentengebrauch nützlich, sollten aber im klinischen Setting
363 separat erhoben werden. Daher wurde auch im HDQ-G auf einen Zusatz zum
364 Medikamentengebrauch verzichtet.

365 An den kognitiven Interviews über die präfinale deutsche Version des HDQ nahmen Personen
366 mit fließenden Deutschkenntnissen teil, die zum Zeitpunkt der Erhebung in der
367 deutschsprachigen Schweiz wohnhaft waren. Der Gebrauch der deutschen Schriftsprache in der
368 Schweiz ist weitgehend identisch mit dem aus anderen deutschsprachigen Ländern, mit wenigen
369 semantischen oder grammatikalischen Unterschieden, welche für diese Studie als nicht relevant
370 erachtet wurden. Daher kann der HDQ-G aus Sicht des Expert*innenkomitees auch in allen
371 anderen deutschsprachigen Ländern ohne Bedenken verwendet werden.

372 Zur Reproduzierbarkeit des HDQ-G wurde eine Test-Retest-Reliabilität bei n=29 stabilen
373 Teilnehmer*innen mit einem sehr guten ICC-Wert von 0.89 und einem 95% KI von 0.77 bis 0.95
374 berechnet. Ein ähnlicher Wert (ICC = 0.93) wurde für die Test-Retest-Reliabilität der
375 englischsprachigen Originalversion des HDQ mit n=61 Proband*innen erzielt (Muller 2007).
376 Allerdings wurden dabei beide Messungen mit einem kürzeren Abstand von nur 24 Stunden
377 durchgeführt (Muller, 2007). In der aktuellen Studie betrug der Abstand im Durchschnitt 4.37
378 (SA=3.30) Tage. Bei der Ursprungsversion des HDQ wurde zusätzlich eine dritte Messung nach
379 30 Tagen mit noch n=18 Teilnehmer*innen durchgeführt. Dabei zeigte sich, mit einem ICC-Wert
380 von 0.90, dass der HDQ die kopfschmerzbedingten Funktionseinschränkungen auch über diesen
381 Zeitraum noch sehr gut abbilden kann, solange die Teilnehmerinnen ihre Situation als
382 „unverändert“ einstufen (Muller 2007). Ähnlich gute bis sehr gute Reliabilitätswerte zeigte auch
383 die deutschsprachige und französischsprachige Version des MIDAS (Benz et al. 2018) (Magnoux
384 et al. 2008). Die geringen Unterschiede könnten mit unterschiedlichen zeitlichen Abständen
385 zwischen den Erhebungen und auch in der Dauer der Kopfschmerzgeschichte liegen, so zeigten
386 sich chronische Kopfschmerzpatient*innen stabiler als Episodische (Magnoux et al. 2008). Eine
387 Richtlinie für einen optimalen Abstand zwischen zwei Messungen besteht nicht, in der
388 vorliegenden Studie wurden die Teilnehmer*innen via E-Mail nach 2-4 Tagen aufgefordert den
389 Fragebogen ein 2-tes Mal auszufüllen. Jedoch reagierten die meisten Teilnehmer*innen erst mit
390 Verzögerung oder auch erst nach nochmaliger Aufforderung durch die Studienleitung. Je nach
391 Akuität oder Chronizität der Teilnehmer*innen, und damit der „Stabilität“ des untersuchten
392 Phänomens, sollten in künftigen Studien auch die „Erinnerungsaufrufe“ geplant werden. In der
393 aktuellen Studie nahmen durchschnittlich viele chronische Kopfschmerzpatient*innen teil, was die
394 sehr gute Reliabilität (mit-)erklären kann (Tabelle 1), dies war jedoch a priori kein
395 Einschlusskriterium, und ist eher dem Zufall zu verdanken. Bei Untersuchungen zur Test-Retest-
396 Reliabilität von vier Fragebögen (u.a. SF-36) bei Personen mit Kniebeschwerden innerhalb eines
397 Zeitraums von zwei Tagen bis zwei Wochen konnten keine signifikanten Differenzen zwischen
398 den Erhebungs-Abständen festgestellt werden (Marx et al, 2003). Die Ergebnisse dieser Studie
399 lassen sich für methodologisch ähnliche Studien zur Test-Retest-Reliabilität verallgemeinern
400 (Marx et al, 2003).

401 Die Mittelwertdifferenz zwischen beiden Messungen lag beim HDQ-G bei +1.72 Punkten, und
402 95% aller Werte lagen innerhalb von +10.77 und -7.33 Punkten. Dies gibt bereits erste Hinweise
403 auf die Veränderungssensitivität oder Responsivität des HDQ-G, diese zu bestimmen war jedoch
404 kein Ziel der vorliegenden Studie. (Giavarina, 2015). In der Originalversion wurden präzisere
405 Werte für die Mittelwertdifferenz mit +0.84 und 95% LoA von +7.65 und -5.97 berechnet, was ca.

406 15% des maximal möglichen Gesamt-Scores ausmacht, im Vergleich zu ca. 20% in dieser Studie
407 (Muller 2007). Ähnliches gilt auch für die deutschsprachige MIDAS-Version (Benz et al. 2018).

408 Der HDQ-G zeigt somit eine gute Test-Retest-Reliabilität mit akzeptablen Grenzen der
409 Übereinstimmung. Ein Einsatz im ambulanten physiotherapeutischen Setting erscheint möglich,
410 sowie auch dessen Anwendbarkeit bei spezifischen Kopfschmerzpopulationen und anderen
411 nicht-medikamentösen Settings. Ein Vergleich des HDQ-G mit anderen, etablierten
412 deutschsprachigen Messinstrumenten wie MIDAS und HIT-6, scheint indiziert.

413 **Stärken und Limitationen**

414 Eine der Stärken dieser Studie ist der professionelle und kurze Übersetzungsprozess, inklusive
415 Einbezug eines Expert*innenkomitee und Original- Autor. Außerdem wurde eine ausreichend
416 große Stichprobe an Teilnehmer*innen für die kognitiven Interviews rekrutiert.

417 Zu den Limitationen der Studie zählt die niedrige Anzahl männlicher Probanden, die an der Studie
418 teilnahmen. Dies sorgte für ein ungleiches Geschlechter-Verhältnis, entspricht aber weitgehend
419 den Prävalenzen für primäre Kopfschmerzen wie Migräne und Spannungskopfschmerzen (Porst
420 et al. 2020; Stovner et al. 2018).

421 Einige Teilnehmer*innen (n=16) brachen die Online-Umfrage ab, nachdem sie aufgefordert
422 wurden, ihre persönlichen Daten anzugeben. Auf eine anonyme Erhebung wurde jedoch
423 zugunsten von rückverfolgbaren und validen Daten verzichtet, außerdem mussten persönliche
424 Daten (mindestens E-Mail-Adresse) erfasst werden, um die Teilnehmerinnen zur zweiten
425 Erhebung einzuladen. Jedoch könnten nur einmalige und ausschließlich Online-Daten
426 Erhebungen auch anonym erfolgen, was die die Drop-Out-Rate wahrscheinlich reduziert.

427 Die bereits erwähnten verzögerten Antworten auf die Einladung zur 2. Umfrage konnten auch
428 darauf zurückgeführt werden, dass die Erinnerungs-emails in Spam Ordner einiger
429 Teilnehmer*innen landeten. Weitere technische Limitationen waren auch etwas verzerrte
430 Darstellungen der Umfragen, je nachdem ob man diese auf PCs, Tablets oder Smartphones
431 öffnete. Vorgängige Testmessungen wurden für diese Studie zwar durchgeführt, konnten aber
432 offensichtlich nicht alle zu erwartenden Probleme simulieren. Ob der Fragebogen für
433 Populationen mit ausschließlich spezifischen Kopfschmerz-Typen wie Migräne oder
434 Spannungskopfschmerzen oder bei Adoleszenten geeignet ist kann, zurzeit nicht beantwortet
435 werden und müsste gesondert untersucht werden.

SCHLUSSFOLGERUNG

Der HDQ-G ist ein für den deutschsprachigen Raum neu validierter Fragebogen für den Einsatz in den erwachsenen deutschsprachigen Populationen im ambulanten klinischen und spezifisch physiotherapeutischen Setting. Er besteht aus neun Fragen zur Schmerzintensität, Aktivitätseinschränkung und Aktivitätsvermeidung von Personen mit Kopfschmerzen. Der HDQ-G zeigt eine gute Test-Retest Reliabilität und akzeptable Übereinstimmung bei stabilen Kopfschmerzpatient*innen innerhalb eines Zeitraums von zwei Tagen bis 2 Wochen zwischen den Messzeitpunkten. Weitere Forschung ist nötig, um seine Eignung als Messinstrument auch außerhalb der Physiotherapie und/oder mit anderen, spezifischeren Kopfschmerz-Typen, und im Vergleich mit anderen Fragebögen zu untersuchen.

ERKLÄRUNGEN

Ethische Aspekte

Die Studie wurde von der kantonalen Ethikkommission des Kantons Zürich genehmigt (ID 2021-01082).

Zustimmung zur Veröffentlichung - Nicht zutreffend

Verfügbarkeit von Daten und Materialien – Nicht zutreffend

Registrierung

Die Studie wurde im Register clinicaltrials.gov registriert (clinicaltrials.gov: NCT04877561).

Finanzielle Unterstützung

Diese Forschung erhielt keine spezifische Finanzierung von öffentlichen, kommerziellen oder gemeinnützigen Stellen.

Interessenkonflikte

Die Autoren geben an, dass kein Interessenkonflikt besteht.

Autor*innen-beiträge

Alle Autoren haben die finale Version des Manuskripts gelesen und genehmigt. Alle Autoren erklären, dass Sie für alle Aspekte der Arbeit verantwortlich sind und sie gewährleisten, dass Fragen im Zusammenhang mit der Richtigkeit oder Integrität eines jeden Teils der Arbeit angemessen untersucht und gelöst werden.

Danksagung

Die Autoren dieser Arbeit bedanken sich bei:

allen Patient*innen welche teilgenommen haben, ausserdem bei Silvia Careddu, Fr. Dr. Andrea Hunziker, Michelle Haas und Dr. Vanessa Klaas von der Zürcher Hochschule für angewandte Wissenschaften (<https://www.zhaw.ch>). Bei Karin Andermatt und der Übersetzungsfirma Traduko®. Bei Rick Peters und den Mitgliedern des SVOMPs (<https://www.svomp.ch>). Bei Ken Niere, dem Autor der originalen englisch-sprachigen Version. Ein weiterer Dank geht an das gesamte Team Physioline (<https://www.physioline.ch>), Ueli Bärtschi und Lena Müller und Kathrin Grötsch. Abschliessend bedanken wir uns bei Gabriela Ivanova, sowie bei Danja Schneider und Detlef Latuski.

REFERENZEN

- Amons, A. L., Castien, R. F., van der Wouden, J. C., De Hertogh, W., Dekker, J., & van der Horst, H. E. (2019). Manual therapy as a prophylactic treatment for migraine: Design of a randomized controlled trial. *Trials*, *20*(1), 785. <https://doi.org/10.1186/s13063-019-3937-8>
- Andersen, C. H., Jensen, R. H., Dalager, T., Zebis, M. K., Sjøgaard, G., & Andersen, L. L. (2017). Effect of resistance training on headache symptoms in adults: Secondary analysis of a RCT. *Musculoskeletal Science and Practice*, *32*, 38–43. <https://doi.org/10.1016/j.msksp.2017.08.003>
- Behr, D. (2017). Assessing the use of back translation: The shortcomings of back translation as a quality testing method. *International Journal of Social Research Methodology*, *20*(6), 573–584. <https://doi.org/10.1080/13645579.2016.1252188>
- Bendtsen, L., & Jensen, R. (2011). *Pharmacological and Non-pharmacological Management of Tension-type Headache*. <https://doi.org/10.17925/ENR.2008.03.01.119>
- Benz, T., Lehmann, S., Gantenbein, A. R., Sandor, P. S., Stewart, W. F., Elfering, A., Aeschlimann, A. G., & Angst, F. (2018). Translation, cross-cultural adaptation and reliability of the German version of the migraine disability assessment (MIDAS) questionnaire. *Health and Quality of Life Outcomes*, *16*(1), 42. <https://doi.org/10.1186/s12955-018-0871-5>
- Bland, J. M., & Altman, D. (1986). STATISTICAL METHODS FOR ASSESSING AGREEMENT BETWEEN TWO METHODS OF CLINICAL MEASUREMENT. *The Lancet*, *327*(8476), 307–310. [https://doi.org/10.1016/S0140-6736\(86\)90837-8](https://doi.org/10.1016/S0140-6736(86)90837-8)
- Deutschl, G., Beghi, E., Fazekas F., Varga, T., Christoforidi K.A., Sipido E.S., Bassetti C.L., Vos T. Feigin V.L, (2020). The burden of neurological diseases in Europe: an analysis for the Global Burden of Disease Study 2017. *The Lancet Public Health*, *5* (10), e551- e567. [https://doi.org/10.1016/S2468-2667\(20\)30190-0](https://doi.org/10.1016/S2468-2667(20)30190-0)

- Bujang, M. A. (2017). A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: A review. *Archives of Orofacial Sciences*, *12*, 1–11.
- Epstein, J., Osborne, R. H., Elsworth, G. R., Beaton, D. E., & Guillemin, F. (2015). Cross-cultural adaptation of the Health Education Impact Questionnaire: Experimental study showed expert committee, not back-translation, added value. *Journal of Clinical Epidemiology*, *68*(4), 360–369.
<https://doi.org/10.1016/j.jclinepi.2013.07.013>
- Giavarina, D. (2015). Understanding Bland Altman analysis. *Biochemia Medica*, *25*(2), 141–151.
<https://doi.org/10.11613/BM.2015.015>
- Haywood, K. L., Mars, T. S., Potter, R., Patel, S., Matharu, M., & Underwood, M. (2018). Assessing the impact of headaches and the outcomes of treatment: A systematic review of patient-reported outcome measures (PROMs). *Cephalalgia*, *38*(7), 1374–1386.
<https://doi.org/10.1177/0333102417731348>
- International Headache Society. (2018). Headache Classification Committee of the International Headache Society (IHS) The International Classification of Headache Disorders, 3rd edition. *Cephalalgia*, *38*(1), 1–211. <https://doi.org/10.1177/0333102417738202>
- Klein, E. (1993). Konditionalsätze und ihre pragmatischen Funktionen im Deutschen, Englischen und Französischen. In J. Darski & Z. Vetulani (Hrsg.), *Sprache—Kommunikation—Informatik* (S. 431–436). DE GRUYTER. <https://doi.org/10.1515/9783111566665.431>
- Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, *15*(2), 155–163.
<https://doi.org/10.1016/j.jcm.2016.02.012>
- Kosinski, M., Bayliss, M. S., Bjorner, J. B., Ware, J. E., Garber, W. H., Batenhorst, A., Cady, R., Dahlöf, C. G. H., Dowson, A., & Tepper, S. (2003). A six-item short-form survey for measuring headache

- impact: The HIT-6™. *Quality of Life Research*, 12(8), 963–974.
<https://doi.org/10.1023/A:1026119331193>
- Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1), 159–174. <https://doi.org/10.2307/2529310>
- Luedtke, K., Basener, A., Bedei, S., Castien, R., Chaibi, A., Falla, D., Fernández-de-las-Peñas, C., Gustafsson, M., Hall, T., Jull, G., Kropp, P., Madsen, B. K., Schaefer, B., Seng, E., Steen, C., Tuchin, P., von Piekartz, H., & Wollesen, B. (2020). Outcome measures for assessing the effectiveness of non-pharmacological interventions in frequent episodic or chronic migraine: A Delphi study. *BMJ Open*, 10(2), e029855. <https://doi.org/10.1136/bmjopen-2019-029855>
- Magnoux, E., Freeman, M. A., & Zlotnik, G. (2008). MIDAS and HIT-6 French translation: Reliability and correlation between tests. *Cephalalgia: An International Journal of Headache*, 28(1), 26–34. <https://doi.org/10.1111/j.1468-2982.2007.01461.x>
- Martin, M., Blaisdell, B., Kwong, J. W., & Bjorner, J. B. (2004). The Short-Form Headache Impact Test (HIT-6) was psychometrically equivalent in nine languages. *Journal of Clinical Epidemiology*, 57(12), 1271–1278. <https://doi.org/10.1016/j.jclinepi.2004.05.004>
- Marx, R. G., Menezes, A., Horovitz, L., Jones, E. C., & Warren, R. F. (2003). A comparison of two time intervals for test-retest reliability of health status instruments. *Journal of Clinical Epidemiology*, 56(8), 730–735. [https://doi.org/10.1016/S0895-4356\(03\)00084-2](https://doi.org/10.1016/S0895-4356(03)00084-2)
- McGraw, H. (2017). Chapter 26—Statistical Measures of Reliability. In L. G. Portney & M. P. Watkins, *Foundations of Clinical Research: Applications to Practice* (3. Aufl.). <https://fadavispt.mhmedical.com/content.aspx?bookid=1987§ionid=149505679#11382528>

- Mohler, P. P., Dorer, B., de Jong, J., & Hu, M. (2016). *Translation: Overview. In Guidelines for Best Practice in Cross-Cultural Surveys (pp. 233-377), MI: Survey Research Center, Institute for Social Research, University of Michigan.* <http://ccsg.isr.umich.edu/>.
- Mokkink, L., Prinsen, C., Patrick, D., Alonso, J., Bouter, L., CW de Vet, H., & Terwee, C. (2019). *COSMIN Study Design checklist for Patient-reported outcome measurement instruments.* <https://www.cosmin.nl/>
- Moore, C. S., Sibbritt, D. W., & Adams, J. (2017). A critical review of manual therapy use for headache disorders: Prevalence, profiles, motivations, communication and self-reported effectiveness. *BMC Neurology, 17.* <https://doi.org/10.1186/s12883-017-0835-0>
- Muller, P. (2007). *Reproducibility and responsiveness of the headache disability questionnaire.* [Honorsthesi]. La Trobe University, Melbourne.
- Muñoz-Gómez, E., Inglés, M., Serra-Añó, P., & Espí-López, G. V. (2021). Effectiveness of a manual therapy protocol based on articulatory techniques in migraine patients. A randomized controlled trial. *Musculoskeletal Science and Practice, 54,* 102386. <https://doi.org/10.1016/j.msksp.2021.102386>
- Niere, K. (2009). Chapter 13—Measurement of headache. In P. Selvaratnam, K. Niere, M. Zuluaga, S. Friedmann, C. Sloan, & E. Byrne (Hrsg.), *Headache, Orofacial Pain and Bruxism* (S. 153–165). Churchill Livingstone. <https://doi.org/10.1016/B978-0-443-10310-0.00013-7>
- Niere, K., & Quin, A. (2009). Development of a headache-specific disability questionnaire for patients attending physiotherapy. *Manual Therapy, 14(1),* 45–51. <https://doi.org/10.1016/j.math.2007.09.006>
- Pohontsch, N., & Meyer, T. (2015). Das kognitive Interview – Ein Instrument zur Entwicklung und Validierung von Erhebungsinstrumenten. *Die Rehabilitation, 54(01),* 53–59. <https://doi.org/10.1055/s-0034-1394443>

Porst, M., Wengler, A., Leddin, J., Neuhauser, H., Katsarava, Z., Von Der Lippe, E., Anton, A., Ziese, T., & Rommel, A. (2020). *Migräne und Spannungskopfschmerz in Deutschland. Prävalenz und Erkrankungsschwere im Rahmen der Krankheitslast-Studie BURDEN 2020*.

<https://doi.org/10.25646/6988.2>

Stewart, W. F., Lipton, R. B., Dowson, A. J., & Sawyer, J. (2001). Development and testing of the Migraine Disability Assessment (MIDAS) Questionnaire to assess headache-related disability. *Neurology*, 56(suppl 1), S20–S28. https://doi.org/10.1212/WNL.56.suppl_1.S20

Stovner, L. J., Nichols, E., Steiner, T. J., Abd-Allah, F., Abdelalim, A., Al-Raddadi, R. M., Ansha, M. G., Barac, A., Bensenor, I. M., Doan, L. P., Edessa, D., Endres, M., Foreman, K. J., Gankpe, F. G., Gopalkrishna, G., Goulart, A. C., Gupta, R., Hankey, G. J., Hay, S. I., ... Murray, C. J. L. (2018). Global, regional, and national burden of migraine and tension-type headache, 1990–2016: A systematic analysis for the Global Burden of Disease Study 2016. *The Lancet Neurology*, 17(11), 954–976. [https://doi.org/10.1016/S1474-4422\(18\)30322-3](https://doi.org/10.1016/S1474-4422(18)30322-3)

Willis, G. B. (2004). *Cognitive interviewing: A tool for improving questionnaire design*.

<https://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=1995048>