The validity and clinical utility of the Disabilities of the Arm Shoulder and Hand (DASH) questionnaire for hand injuries in developing country contexts: A Systematic Review

Abstract:

Study Design: Systematic review.

Introduction: The purpose of this study was to systematically review the evidence available on the validity and clinical utility of the DASH as a measure of activity and participation in patients with musculoskeletal hand injuries in developing country contexts.

Methods: We conducted a PROSPERO-registered comprehensive literature search and extracted descriptive data. Two reviewers independently assessed methodological quality with the COlSensus-based Standards for the selection of health Measurement INstruments (COSMIN) critical appraisal tool, the checklist to operationalize measurement characteristics of patient rated outcome measures and the multi-dimensional model of clinical utility.

Results: Fourteen studies reporting 12 language versions met the eligibility criteria. Two language versions (Persian and Turkish) had an overall rating of good, and one (Thai) had an overall rating of excellent for cross-cultural validity. The remaining nine language versions had an overall poor rating for cross-cultural validity. Content and construct validity and clinical utility yielded similar results.

Discussion/Conclusions: Poor quality ratings for validity and clinical utility were due to insufficient documentation of results and inadequate psychometric testing. With the increase in migration and globalization, hand therapists are likely to require a range of culturally adapted and translated versions of the DASH. Recommendations include rigorous application and reporting of cross-cultural adaptation, appropriate psychometric testing and testing of clinical utility in routine clinical practice.

Level of Evidence: 2c

Keywords: Disabilities of the Arm, Shoulder and Hand questionnaire, validity, utility, hand injury
Introduction

The Disabilities of the Arm Shoulder and Hand (DASH) Questionnaire is an extensively researched evaluative and discriminative region specific patient rated outcome measure (PROM) used by many clinicians and researchers in the field of hand therapy.\(^1\) This instrument was first developed by the American Academy of Orthopedic Surgeons, the Council of the Musculoskeletal Speciality Societies and the Institute for Work and Health (IWH), Toronto (Ontario) and published in 1996 by Hudak et al.\(^1\) The DASH measures symptoms, and some aspects of activity and participation according to the nine domains outlined in the International Classification of Functioning, Disability and Health (ICF) in patients with musculoskeletal conditions of the upper limb.\(^1-4\) Assessments of upper extremity function used in routine hand therapy practice have traditionally focused on aspects of body function and structure (such as the measurement of range of motion or strength) which are clinician derived rather than patient reported.\(^5\) In more recent publications a number of authors have investigated the advances in the use of instruments addressing aspects of activity and participation in addition to the predictable use of instruments that measure a single dimension such as strength or sensation.\(^5-8\) The implementation of and call for more client-centered approaches, addressing the broader understanding of health brought about by adopting the ICF framework, which also encompasses a patient perspective, has paved the way for greater use of PROMs that assess aspects of activity and participation.\(^9\)

In South Africa, the routine use of measures of activity and participation remains low.\(^10\) Therapists offer time constraints and lack of applicability in the practice context as reasons for non-use of the DASH.\(^10\) Time constraints are a common reason for non-use of PROMs.\(^6,7,11\) In contrast, the quick administration time of the DASH has been reported in some studies.\(^12,13\) It is however worth exploring the notion of lack of applicability and time constraints associated with using such a rigorous instrument, with well-established psychometric properties, in this context. A systematic review of the cross-cultural adaptation of the DASH included only English language publications (n = 9); eight of nine from developed country contexts.\(^13\) This presents a biased view in research on this topic for developed countries.

Diverse cultures, languages and occupations make providing interventions in developing contexts more complex. Contextual variation and diversity culminates in differences in the execution and experience of daily activities, occupations and the type of occupations performed. In client-centered care, these differences (essentially in activity and participation) have to be captured, considered and appreciated in daily encounters with patients. Using PROMs is one way to do this. Alotaibi states that the "availability [and use] of assessments [that were] adapted for use in a different culture promotes the
client’s capacity to engage in culturally meaningful occupations.” It is therefore essential to evaluate whether a measure such as the DASH measures the constructs it appears to measure in patients with hand injuries in a developing country context.

The Consensus-based Standards for the Selection of Health Measurement Instruments (COSMIN) checklist was devised to assist researchers and clinicians to evaluate the psychometric and clinimetric properties of health related measurement instruments. It defines the measurement properties that should be assessed, and the criteria for acceptable measurement. COSMIN defines cross-cultural validity as “The degree to which the performance of the items on a translated or culturally adapted HR-PRO instrument are an adequate reflection of the performance of the items of the original version of the HR-PRO instrument.” Content validity is the relevance of the items of the measurement instrument to the construct of interest, and construct validity refers to the ability of an instrument to measure the theoretically intended constructs. In accordance with the COSMIN criteria construct validity is evaluated by considering structural validity (through factor analysis), hypothesis testing (through moderate correlations with instruments measuring the same construct) and cross-cultural validity (by evaluating differences in factor structure or differential item function (DIF) between language versions). Francis et al incorporated knowledge from the COSMIN criteria and presented a simplified checklist for evaluating the methodological quality of PROMs. They concluded that their checklist could assist researchers or clinicians with varied expertise and experience in measurement theory to evaluate the quality of the PROM in systematic reviews or for use in clinical practice. Francis et al included responsiveness (longitudinal construct validity) and predictive validity as a form of criterion-related validity. In the present review we considered cross-cultural, construct and content validity.

A further consideration was the clinical utility of the DASH. The complexity of clinical utility makes its evaluation a challenge. Clinical utility is defined as the usefulness of an assessment or intervention in clinical practice. The usefulness of the DASH cannot be contested; this is clear from the multiple language versions and extensive use of the measure in clinical practice and research. In addition, the DASH can be used to assess the functional status of traumatic hand injured patients. However, therapists in a developing context do not find the DASH useful due to lack of applicability. In accordance with Smarts’ conceptualization of clinical utility, therapists may not have found the instrument to benefit their treatment approach, or the patient. Smart summarized the dimensions of clinical utility and identified the components to be appropriate, accessible, practicable and acceptable. Corr and Siddons highlighted the validations of the measure for the relevant client group to be an important consideration for clinical utility. Information on the clinical utility of the DASH as a measure
of activity and participation in patients with hand injuries in developing country contexts is imperative to make decisions about using it for its intended purpose.

**Purpose of the review**

The purpose of this systematic review was to examine the validity and clinical utility of the DASH questionnaire as a measure of activity and participation in patients with hand injuries in a developing country context.

**Methods**

**Search strategy**

We conducted a PROSPERO-registered comprehensive literature search using the following key electronic databases: MEDLINE (PubMed), EBSCOHost (Academic Search Premier, CINAHL, and Africa Wide), Scopus, Web of Science and Google Scholar. We searched grey literature was searched through the World Health Organization Library OpenGrey and OpenDOAR. Search terms included: Disabilities of the Arm, Shoulder and Hand questionnaire, cross-cultural adaptation, validity, clinical utility and musculoskeletal hand injury. See Supplementary file 1 (available online) for the electronic database search strategy. Covidence (https://www.covidence.org) was used to manage the review. The first author completed the database searches, scanned for relevance based on the title and abstract and applied the inclusion criteria. The first and second author applied all eligibility criteria against the full text of the remaining articles to select relevant studies for the review. The first author reviewed reference lists of relevant articles and performed hand searches to identify all appropriate studies. There was agreement among the authors as to which articles to include in this systematic review.

**Identification and selection of studies**

Inclusion criteria were any studies of the DASH questionnaire from inception to 2016, all languages, with a study population of adults (age ≥ 18) with musculoskeletal (MSK) hand injury, and from developing country contexts*. Study aims had to include evaluation of, or reporting on, validity and/or clinical utility. We excluded trials that used the DASH as an outcome measure without studying the measurement properties in question.

**Data extraction and assessment of methodological quality**

* Developing country context is understood to be middle income (upper and lower) and low-income countries according to the World Bank Rankings.23
We gathered descriptive information (such as, country, income group, language and study design) on included studies (n=14) (Figure 1; Table 1). Two reviewers independently assessed the methodological quality of the selected studies (Xxx & XX). Cross-cultural validity was assessed with the COSMIN checklist (http://www.cosmin.nl/cosmin_checklist.html).\textsuperscript{15-18} The COSMIN group advocates using this checklist as a modular tool as the sections are not summarized as total scores. The scoring system comprises a 4-point rating scale (excellent, good, fair or poor) to rate each measurement property. We used the cross-cultural validity section to calculate the quality score.\textsuperscript{14-18} The checklist to operationalize measurement characteristics of PROMs and the Multi-dimensional model of clinical utility were used to assess content and construct validity and clinical utility (respondent burden and presentation, appropriateness and acceptability).\textsuperscript{16,19,20} We used a dichotomised scoring system to rate these properties in each publication. Differences in ratings were discussed and resolved through consensus. A third reviewer (XXX) was available to independently appraise these articles, this was however not required.

**Synthesis of results**

The 15 cross-cultural validity properties from the COSMIN checklist were scored according to the COSMIN guidelines’ “worst score counts” method, i.e. if one item per box was scored “poor” then the overall score for that particular measurement property was poor.\textsuperscript{14,16} Four properties for construct validity and three properties for content validity were scored by awarding a “yes” if the property was addressed, “no” if not addressed and “not applicable” if that form of validity was not an aim of the study. The same method was applied for the properties concerning clinical utility namely: respondent burden and presentation (3 properties), appropriate (3 properties) and acceptable (2 properties). We derived an overall account of validity and clinical utility from the data.

**Results**

Table 1 summarize characteristics of included studies.\textsuperscript{24-37} The 14 articles reported 12 language versions from 11 low and middle-income countries (Figure 2). Thirteen studies were in English and one in Turkish. The Turkish study was translated into English for analysis. There were ten published articles and one of each of the following: congress poster presentation research report, letter to a journal editor, newsletter article and conference presentation. Language versions were from five upper middle-income countries (Brazil, Malaysia, Iran, Thailand and Turkey), five lower middle-income countries (Armenia, India, Nigeria, Russia and Sri Lanka) and one low-income country (Ethiopia). Two studies reported on languages from the African continent (Nigeria and Ethiopia) and there were two separate
Thai versions (the KKU-DASH and the DASH-TH). Both the Brazilian and Turkish DASH had two studies reporting psychometric properties. The median number of study participants was 40 (range: 30 to 309). Studies included male and female participants, apart from the Malay study that had only male participants. There were a variety of settings including outpatient clinics and orthopedic inpatient services addressing MSK hand injury. Sampling techniques were often not reported.

**Cross-cultural validity**

Cross-cultural validity was excellent for one of the Thai versions (DASH-TH) and good for the Turkish and Persian versions (Table 2). All other language versions scored poor for cross-cultural validity, even though the Nigerian (Yoroba) and Brazilian Portuguese versions had excellent scores for 11 and 10 of the 15 properties respectively (Table 2). When permission was granted for translation of the instrument, the researchers employed Beaton et al.’s recommendations for cross-cultural adaptation. Reporting was not clear in all instances, for example, whether translators worked independently, if translations were reviewed by a committee or whether the instrument was field tested to check the interpretation, cultural relevance and ease of comprehension. The changes made to the different language versions towards addressing cultural relevance were of interest. The DASH questionnaire item number one, “Opening a tight or new jar”, was culturally adapted to a “well corked bottle” in Nigerian Yorobu and “a threaded-lid tight or new jar” in Russian and Armenian. Item seven, “Do a heavy household chore (e.g. wash walls, wash floors)”, was adapted to “fetching water from the well” in Nigerian Yorobu. Washing walls or floors is not considered a heavy task in Nigerian culture, but fetching water from a well is undertaken often and considered “heavy”. The DASH–TH (Thai) version adapted item 12 “Change a light bulb overhead” to “sweeping the ceiling” as few people in Thailand change a light bulb, but they are likely to sweep the ceiling (requiring the same shoulder, arm and hand range and ability). The unit of weight measurement was changed to match the local measurement system, and items 18 and 19 (recreational activities) were changed to activities undertaken within that context. The study reporting on the Amharic version (Ethiopia) had no reported data on cross-cultural validity.

**Content and construct validity**

Content validity was not reported in 57% (8 of 14) of studies. Research on the Brazilian Portuguese, Persian, Nigerian and Turkish versions addressed all three properties of content validity. In most studies (7 of 8) content experts were involved in developing the new language version. In terms of construct validity, 85% (12 of 14) of studies evaluated the correlation between the new language version and the existing English language version or other relevant data. Factor analysis was employed in the Nigerian,
204 Persian and Brazilian Portuguese language versions of the DASH. Principal component analysis was
205 performed on the Persian and Nigerian DASH with the Persian meeting the assumptions of
206 unidimensionality and the Nigerian justifying seven subscales. For the Brazilian Portuguese DASH
207 exploratory factor analysis was performed and provided justification for three subscales. Content and
208 construct validity are reported Table 3.

209 Clinical utility

210 Clinical utility was evaluated on three concepts namely respondent burden and presentation,
211 appropriateness and acceptability. Respondent burden and presentation contained three criteria of
212 which time taken to complete was reported in 64% (9 of 14) of the studies. Seven of 14 studies reported
213 the required literacy level to complete the questionnaire. Appropriateness (importance of the measure
214 to clinical decision making) was reported in all 14 studies. The acceptability of the measure to the client,
215 family or carers was reported in 57% (8 of 14) of the studies. Acceptability to the society as a whole or to
216 stakeholder groups was not addressed in any studies. Refer to Table 3 for an overall account of clinical
217 utility criteria.

218 Discussion

219 This systematic review aimed to examine the cross cultural, construct and content validity, and
220 clinical utility of the DASH questionnaire as a measure of activity and participation in patients with hand
221 injuries in developing country contexts. Guillemin et al 39 proposed guidelines for the cross-cultural
222 adaptation of health related quality of life measures that were used to develop the recommendations
223 for the cross-cultural adaptation of the DASH and the QuickDASH. 38 Guillemin et al offered instances
224 under which measures have to undergo translation and adaptation. 39 Beaton et al expanded these by
225 recommending that when the DASH is used in another language and country, the resulting changes in
226 language and culture requires the instrument to be translated and culturally adapted. 38,39 The cross-
227 cultural equivalence of the source and final versions must include aspects of semantic, idiomatic,
228 experiential and conceptual equivalence. 39 Conceptual equivalence should address aspects of cultural
229 beliefs understood to be “who people are, how they interact with the world and how they behave in
230 certain situations”. 40 p.1 The advantages of an outcome measure such as the DASH, that has undergone
231 rigorous methodical implementation of cross-cultural adaptation to produce equivalent versions,
232 irrespective of language, country or culture, are indeterminable. Guillemin et al explains these
233 advantages to include the possibility of making comparisons between countries and cultural groups and
234 involving less cost and time than developing a new measure. 39 The process of establishing a common
measure for the investigation of the construct(s) in different cultural contexts does not end with the
cross-cultural adaptation of the measure. The psychometric properties of the adapted measure have to
be verified through further testing. Clinical utility is an important property to consider with the same
rigor, as the DASH only holds value if it is used.

The process of cross-cultural translation is assumed to ensure the retention of aspects of validity
and reliability. However, Beaton et al and Guillemin et al recognize that due to the differences in
cultures, items may be rated as less or more difficult in the adapted versions, and would therefore
require psychometric testing of the new language version. The example from the Yoruba version of
the DASH, “fetching water from the well” to measure the execution of heavy household chores, as
opposed to “washing walls or floors”, illustrates this point. A number of studies in this review did
address construct validity (8 of 12 language versions across 12 of 14 studies). The instruments used in
this process of hypothesis testing included cross-culturally adapted versions of the SF-36, numeric pain
rating scales and the English version of the DASH. A cross-culturally adapted version of the SF-36 may
not be available for this purpose in some developing contexts. Furthermore, in low socio-economic
populations with low literacy levels, completing the English DASH (often a third language) is also not a
viable option. Numeric rating scales used for this purpose have had mixed results. Establishing
construct validity of a cross-culturally adapted version of the DASH in a developing country is a
challenging prospect.

The Translation and Cultural Adaptation (TCA) group formed by the International Society for
Pharmacoeconomics and Outcomes Research (ISPOR) systematically considered approaches to
translation and cross-cultural adaptation of PROMs. During their review of existing approaches they
noted that the lack of consistency in terminology and methodology resulted in new language versions
with uncertain cross-cultural validity. The recommendations for cross-cultural adaptation available
from the IWH are clear about the process towards establishing an adapted version of the instrument for
use in other countries and cultures. It is interesting to note that despite the clarity of this process, the
analysis of the methodological quality of the cross-cultural validity of the DASH translations included in
our review, yielded mostly poor results. Two possible explanations for this are: 1) the process was not
duly performed, or 2) the process was not duly reported. The COSMIN criteria used to analyze cross-
cultural validity are considered extremely rigorous for evaluating the methodological quality of PROMs.
However, the ten step translational and cross-cultural adaptation ISPOR TCA Principles of Good Practice
guideline highlights additional aspects to be performed and reported that are not considered in the
COSMIN criteria for cross-cultural validity. Evaluating the 12 language versions against the cross-cultural
adoption process proposed by the IWH may demonstrate that due processes were completed and that
the language versions are sufficiently adapted to undergo further psychometric testing. A 2008
systematic review on the cross-cultural adaptation of the DASH concluded that researchers could not
benefit from the experience of cross-cultural adaptation as many articles failed to report details of the
process.\textsuperscript{13} Alotaibi further recommended that subsequent publications address issues encountered
during the process.\textsuperscript{13} However, we did not find documentation on the process for some language
versions from developing contexts available from the DASH website. Certainty about the cross-cultural
validity of those versions therefore eludes us.

Patient literacy levels in developing countries may also impact the use of self-report measures. The
short time taken to complete the DASH is an advantage and may result in greater clinical utility in busy
clinical settings.\textsuperscript{1} In both studies investigating the Brazilian Portuguese version of the DASH, the
researchers completed the questionnaire by interview due to low literacy levels of participants. The time
taken to complete was therefore more than the usual ten minutes. It is not clear how this influences the
use of the measure in routine clinical practice. Arguments exist that PROMs administered by interview
yield better results that when self-completed.\textsuperscript{42} In developing contexts and even in developed contexts
where migration and globalization are a reality, therapists may often not be able to speak the language
spoken by the patient. Completing the DASH by interview brings additional challenges as using
translators not formally trained could change what is communicated to patients to fit the cultural
context.\textsuperscript{40} In the case of the DASH, this practice could alter the validity of the questionnaire. The process
of cross-cultural adaptation followed by additional psychometric testing of the questionnaire should
however prevent this.

Exploring the acceptability and appropriateness of a measure is essential in evaluating its clinical
utility. In some cultures, patients perceive self-reporting to be indicative of healthcare professionals not
knowing what the problem is.\textsuperscript{43} Therapists should see the value of the measure with regards to clinical
decision making (appropriate) and patients should see the measure as valuable in adding to their care
(acceptable). It is important to consider the cultural nuances and differences in the way patients account
for difficulties experienced, through the careful consideration of all steps in the process of cross-cultural
adaptation and through responsible documentation thereof. This could result in therapists finding the
measure appropriate and patients experiencing it as acceptable.

Strengths of this review include using the COSMIN checklist to assess methodological quality
(specifically in relation to cross-cultural validity), PROSPERO registration, and inclusion of any language.
As the COSMIN criteria and checklist were published in 2010, included articles published prior to this are unlikely to have considered them in establishing cross-cultural validity. An additional limitation is the use of the “worst score count” method, which makes a broad overall judgment on the properties considered. Furthermore, the quality assessments used for rating the content and construct validity and clinical utility, have not previously been used in systematic reviews of this nature; therefore, results from these assessments should be considered with caution.

**Conclusion**

The overall validity of the measures ranged between good and poor with only one yielding excellent results. Reasons for this could be inadequate documentation of cross-cultural validation and/or psychometric testing. The usefulness of the DASH in developing contexts is unclear as clinical utility was not routinely reported. While this systematic review focuses on developing contexts the results have significance to the broader readership of this journal. With the increase of migration in addition to globalization, clinicians may encounter patients from other cultures, speaking other languages in any hand therapy setting around the world. Recommendations for practice and research include greater collaboration across cultures, continents and language groups towards: 1) ensuring rigorous application of cross-cultural adaptation and appropriate psychometric testing and; 2) reporting clinical utility in routine clinical practice.

**Figure 1:** PRISMA Flow diagram for study selection

**Table 1:** Characteristics of the included studies

NR = Not reported, UMI = Upper middle income, LMI = Lower middle income, LI = Low income

**Figure 2:** Countries included in the review

**Table 2:** Overall account of cross-cultural validity

Rating: Excellent = ☀️; Good and/or Fair = ☀️; Poor = ☑️; NA = Not applicable; U = Unclear

**Table 3:** Overall account of validity and clinical utility

Rating: √ = Yes, × = No, NA = Not applicable
References:


