Evaluation of shoulder disability questionnaires used for the assessment of shoulder disability after neck dissection for head and neck cancer

David P. Goldstein, MD,[1]* Jolie Ringash, MD,[2] Eric Bissada, MD,[1] Yves Jaquet, MD,[1] Jonathan Irish, MD,[1] Douglas Chepeha, MD,[3] Aileen M. Davis, PhD[4]

[1]Department of Surgical Oncology, Princess Margaret Hospital, Department of Otolaryngology–Head and Neck Surgery, University of Toronto, Toronto, Ontario, Canada, [2]Department of Otolaryngology–Head and Neck Surgery, University of Michigan, Ann Arbor, Michigan, [3]Department of Radiation Oncology, Princess Margaret Hospital, University of Toronto, Toronto, Ontario, Canada, [4]Division of Health Care and Outcomes Research, Toronto Western Research Institute, University Health Network, University of Toronto, Toronto, Ontario, Canada.

ABSTRACT: *Background.* Several questionnaires have been used to evaluate shoulder disability after neck dissection. The purpose of this study was to review these measures and highlight their strengths and weaknesses.

Methods. A literature review was performed to identify measures of shoulder disability after head and neck cancer surgery. These measures were evaluated in terms of their methods of development and assessment of their psychometric properties.

Results. Seven questionnaires were identified. Several of the other questionnaires have been well developed but have not had their psychometric properties assessed in the head and neck cancer population. Each questionnaire has its strengths and weaknesses.

Conclusion. The strengths and weaknesses of the shoulder disability questionnaires should be considered when deciding which questionnaire to use. Efforts should be focused on using well-designed questionnaires that have been assessed in this patient population rather than developing or using other questionnaires. © 2013 Wiley Periodicals, Inc. *Head Neck* **36:** 1453–1458, 2014

KEY WORDS: shoulder impairments, shoulder disability, shoulder scales, Disabilities of the Arm, Shoulder, and Hand Questionnaire (DASH), Neck Dissection Impairment Index (NDII), Shoulder Pain and Disability Index (SPADI), American Shoulder and Elbow Surgeons (ASES)

INTRODUCTION

There has been substantial research assessing shoulder morbidity after neck dissection. On review of the literature, it is apparent that there is no consensus on which outcome measure to use in studies assessing shoulder disability after neck dissection. The lack of a standard questionnaire makes the comparison of results between studies difficult. The head and neck cancer literature is composed of numerous studies that have reported shoulder outcomes using data from questionnaires that were designed by the investigators without consideration of accepted principles of questionnaire development.[1] Although other studies have used questionnaires with acceptable methodological development, many of these questionnaires were developed for evaluation of other pathologies and diagnoses and have not undergone assessment of their psychometric properties in the head and neck cancer population.[1] Patients with head and neck cancer have unique problems that require investigations specific to their disease-related issues.[2]

Choosing which instrument to use poses a potential challenge for investigators of shoulder disability after neck dissection. Therefore, we sought to identify the shoulder disability questionnaires that have been used in the head and neck literature and review each instrument in terms of questionnaire development, testing, and assessment of psychometric properties.

MATERIALS AND METHODS

A literature review was performed using Ovid Medline and Embase databases (from 1980 to July 2011) to identify patient-reported outcome (PRO) questionnaires used in studies assessing shoulder disability after neck dissection for head and neck cancer. The electronic search was restricted to articles published in the English language using the following medical subject heading terms or text words: shoulder, upper extremity, disability, activity limitations, impairment, questionnaire, accessory nerve, shoulder syndrome, morbidity, disability, quality of life, neck dissection, and head and neck cancer. The search was supplemented by cross-referencing potentially relevant citations of the identified articles. A list of shoulder disability questionnaires used as outcome measures in these studies was generated. Questionnaires included for review must have met the following criteria: (1) designed to measure the construct of physical symptoms, activity limitations, and/or disability of the shoulder or upper

*Corresponding author: D. P Goldstein, Princess Margaret Hospital, Room 3-952, 610 University Avenue, Toronto, Ontario, M5G 2M9 Canada. E-mail: david.goldstein@uhn.on.ca

extremity, and (2) undergone some recognized standards of questionnaire development (ie, item generation and reduction, and pretesting) and assessment of one or more psychometric properties in any patient population.

Each questionnaire was appraised according to consensus-based standards for the selection of health status measurement instruments (COSMIN) standards, including questionnaire development, ease of use and scoring, and psychometric properties assessed.[3] The psychometric properties assessed in this review included internal consistency, test–retest reliability, measurement error (standard error of the measure [SEM], and/or minimal detectable change [MDC]), content validity, convergent and divergent validity, discriminant validity, and responsiveness. For the purposes of this review, we chose to provide an overview of the properties that are most commonly reported.

RESULTS

There were a total of 307 references and 84 full-text articles reviewed. Forty-three studies that used a PRO measure to assess shoulder disability after neck dissection for head and neck cancer were identified. In 15 studies, the questionnaires were designed by the investigators solely for their study without using recognized standards of questionnaire development and evaluation, and were therefore excluded. Five used pain scales only. Of the remaining 23 studies, the following shoulder disability questionnaires were identified: the Neck Dissection Impairment Index (NDII), the Shoulder Pain and Disability Index (SPADI), the Shoulder Disability Questionnaire (SDQ), Constant's Shoulder Score (CSS), the modified American Shoulder and Elbow Surgeons (ASES) standardized form, Disabilities of the Arm, Shoulder, and Hand questionnaire (DASH), and the Simple Shoulder Test (SST).

Neck Dissection Impairment Index

The NDII is a 10-item questionnaire that was designed to assess quality of life after neck dissection.[4] The items include questions on shoulder and neck-related pain and stiffness, problems with self-care activities, and participation in social activities, recreational activities, or work. The questions related to activities of daily living, work, and recreation, are broad categories and do not ask about specific activities, although examples are provided. The recall period for these items is 4 weeks. For each question, there is a 5-point response option with $0 =$ worst and $5 =$ the best. A total score is converted to a score out of 100, with lower scores representing greater disability. The NDII can be completed in less than 5 minutes and is easy to score.

Item selection was initially performed by interviewing 40 patients who had undergone either a selective neck dissection or modified radical neck dissection. A panel of surgeons, physiotherapists (PTs), and survey specialists performed further item selection and reduction. Fifteen items were generated, which were then pilot tested on a separate group of patients. Items were reduced to the final 10 questions by removing items with retest correlations <0.50 and through exploratory factor analysis. Content validity was determined through patient interviews, expert review, and pilot testing.[4]

For internal consistency, Cronbach's alpha was 0.94.[4] For test–retest reliability, Spearman's correlation for the total score was 0.91 and the individual item correlations ranged from 0.41 to 1.00.[4] There has been no reporting of either an MDC score or SEM. On convergent validity testing, the total score correlation (the statistical measures used were not described) with CSS was 0.85, and the correlation with the domains on the Short Form Health Survey (SF-36) were: role-physical domain ($r = 0.60$); physical functioning domain ($r = 0.50$); bodily pain domain ($r = 0.32$); social functioning domain ($r = 0.62$); and mental health domain ($r = 0.56$).[4] Assessment of divergent validity was not specifically described. Discriminant validity was determined by demonstrating that the NDII was able to discriminate between varying degrees of "shoulder disability" between selective neck dissection and modified radical neck dissection.[4]

The NDII is the only PRO measure specifically developed to assess disability after neck dissection. There are a few potential limitations of the NDII. Patients included in the reliability and validity analyses all had nerve-sparing neck dissections and were not completely representative of the range of severity of shoulder morbidity seen in the population in whom this instrument would be used. The items regarding activity limitations related to activities of daily living, work, and recreation are broad categories that may create some difficulty with the response in patients who have trouble with only some of the activities listed among the examples. The questionnaire includes both morbidity related to the shoulder and the neck and, therefore, may be more applicable to studies assessing overall morbidity after neck dissection and less applicable to those studies primarily focused on shoulder morbidity only. Test–retest reliability was only assessed using Spearman's correlation. There has been no reporting of intraclass correlation coefficients (ICCs). Spearman's correlation only assesses how scores vary together (association) but does not assess the level of agreement. Although the test and retest scores may correlate well, they may not agree, and therefore can give a false sense of reliability of the measure. For most research and clinical applications, the essence of reliability is in agreement between the 2 sets of data, as assessed with ICC.[5] There has been no testing or reporting of responsiveness or cross-cultural evaluation. Distinct cutoff points to reflect severity have not been determined and there is no published population normative data.

Shoulder Pain and Disability Index

The SPADI was developed to measure "pain and disability" in patients with shoulder pain of musculoskeletal, neurogenic, or undetermined origin.[6] The SPADI consists of 13 items in 2 domains: pain symptoms (5 items) and disability/physical function (8 items).[6,7] The recall period for these items is 1 week. The instrument is specific to those with painful shoulder of any etiology.[6–8] The questionnaire is easy to understand and can be completed in approximately 7 minutes.[9,10] A 10 cm visual analog scale (VAS) is used to rate each item. For the pain scale, $0 =$ no pain and $10 =$ worst pain imaginable. For the disability scale, $0 =$ no difficulty and $10 =$ so difficult it requires help. Patients mark activities they do not perform

as not applicable (N/A), which are then excluded from the score. The individual items scores are converted to a 0 to 100 score, with higher scores representing greater pain and disability.[9] Scores can be calculated for each domain or as a total score, which is determined by averaging the scores for the 2 domains of pain and disability.[11]

Item generation and reduction was performed with the input of physicians and PTs and tested in patients with a complaint of shoulder pain attending an ambulatory care clinic.[6] There was no direct input by patients on item development. Items were eliminated based on low test–retest reliability scores and low correlation with shoulder range of motion (ROM).[8] Factor analysis was performed, which supported the subscales of pain and disability.[8] Content validity was determined through expert review, pilot testing, and later with patient input.[12,13]

For internal consistency, Cronbach's alpha ranges from 0.86 to 0.96.[9,14] The ICCs for test–retest reliability range between 0.66 and 0.95.[6–9,15] MDC scores and SEM have been defined.[10,16] The SPADI has been shown to correlate well (Pearson's or Spearman's correlation ≥ 0.70) with other shoulder disability questionnaires (ie, ASES, DASH, and SST).[7,9,12] For discriminant validity, weak correlations were found on correlation with dissimilar measures, such as the SF-36.[7,9,12] The SPADI has also been demonstrated on a limited basis to be able to differentiate between groups.[9] Responsiveness evaluation has also been performed, demonstrating its ability to detect change in patients with musculoskeletal (MSK) disorders undergoing surgical and nonsurgical treatments. Cross-cultural evaluation has been performed in a few languages.[14,17,18]

There are some potential limitations. Item reduction was done based on correlating items with functional assessment, which may have resulted in eliminating important and highly responsive items, particularly since shoulder ROM has been shown to correlate only modestly with patient's estimation of their subjective functioning.[8] The focus of the SPADI is on shoulder pain, which is only one of many potential symptoms or limitations after neck dissection. In addition, the extent of pain is highly variable and it is less prevalent in nerve-sparing neck dissections compared with nerve-sacrificing neck dissections.[1] The benefit of a questionnaire with a focus on pain would depend upon the research question. From a practical standpoint, the scoring system is more time-consuming to perform; however, a modified version with numerical scoring has been developed. There are no normative data or cutoff values to reflect severity.[19] There has been very limited assessment of internal consistency, test-retest reliability in patients with head and neck cancer, and no assessment of construct validity.

Shoulder Disability Questionnaire

The SDQ is a 16-item PRO measure that was developed for the assessment of pain and its impact on the functional limitations of the shoulder in patients with soft tissue disorders of the shoulder.[9,20] Thirteen questions relate to pain with certain upper extremity activities, and 3 questions deal with sleeping, the need to rub the shoulder, and irritability related to shoulder pain.[9] The recall period for these items is 24 hours. The SDQ takes a short time to complete (<5 minutes). Scoring is easy and based on a "yes," "no," or "N/A" response for activity items not performed. A score of 1 is assigned to a "yes" and a score of 0 to a "no" answer. The (total) score is calculated by dividing the number of positively scored items by the total of applicable/completed items and multiplying by 100. A higher score represents greater disability.

Items were selected based on the "functional status limitations" most frequently reported by patients with shoulder disorders and those judged to be important in the evaluation of treatment outcome. Pretesting was performed in patients seen in a PT practice. PTs and researchers, without physician or patient involvement, generated and reduced the list of items.[20] There has been some demonstration of content validity; however, in other reviews of shoulder disability, content validity for the SDQ was rated as doubtful in comparison with other shoulder disability measures.[15,19] Data about factor analysis have not been published.[19]

For internal consistency, Cronbach's alpha ranges from 0.76 to 0.79.[19] There has been only a single study that assessed test–retest reliability with a Spearman's correlation of 0.88 reported.[13,19] SEM or MDC scores have not been reported.[13] For convergent validity, the SDQ correlated poorly with the SPADI ($r = 0.33$)[19] and there has been no assessment of divergent validity. The SDQ has been shown to be able to detect differences in levels of severity of shoulder complaints in primary care patients, however, the ability to detect differences in shoulder complaints was less optimal in a secondary care population.[21] The developers of the SDQ focused on assessment of its responsiveness, demonstrating that it is able to differentiate between self-rated clinically stable and improved subjects as part of their randomized trial.[9,20,21] The ability to detect change has also been reported in studies on treatments for adhesive capsulitis and rotator cuff tendinitis.[19]

In addition to the limitations in questionnaire development and assessment of psychometric properties, it has not undergone any formal assessment of the psychometric properties in patients with head and neck cancers who have undergone neck dissections. Similar to the SPADI, the SDQ has a major focus on pain with the influence of pain on activity limitations rather than the activity limitation itself.[9] The limited response options ("yes," "no," or "N/A") may potentially limit the sensitivity of detecting change or differences between groups. The instrument assesses the ability to perform the activity in the preceding 24 hours. Therefore, important morbidity may not be recorded if that particular activity was not performed in that time period. There are no distinct cutoffs to reflect severity reported and no normative data or cross-cultural evaluation published.

Constant's Shoulder Score

CSS combines active ROM and strength testing with items for pain (1 item) and activity limitations (4 items). The recall period for the activity items is 1 week. It can be completed in 5 to 7 minutes.[19] Pain is scored on VAS with 0 = maximal pain and 15 = no pain. Activity

items are scored on a Likert scale with 0 = worst and 5 = best. Points are also assigned for the ability to complete specific shoulder ROM and strength maneuvers. The subjective assessment consists of a maximum of 35 points and the physical assessment a maximum of 65 points. A formula for calculating a total score out of 100 is described with 0 = worst and 100 = best function.

Methodology on item selection, item reduction, assessment of content validity, and assignment of weights was not described by the developers.[8] For internal consistency, Cronbach's alpha has been reported at 0.37 and 0.60.[19] Test–retest reliability has been formally evaluated in only a few studies, with ICCs of 0.80 and 0.96 reported.[19,22] There is no description of MDC scores or SEM. There was no initial evaluation of construct validity by the developers.[11] Convergent validity was demonstrated by moderate to high correlations (0.49 to 0.87) between CSS and other shoulder-specific questionnaires.[19] Assessment of divergent validity could not be identified. Normative data are available.

Despite the limitations in questionnaire design and only a few studies assessing its psychometric properties, CSS has been extensively used in the MSK literature.[23,24] It has not undergone any assessment of psychometric properties in patients after neck dissection. It has not undergone any cross-cultural adaptation. Although CSS does contain a component asking about shoulder disability, the instrument is weighted heavily on function and pain rather than activity limitations.[8,25] It is also the author's opinion that physical function and subjective assessment of symptoms and activity limitations are best assessed separately.

American Shoulder and Elbow Surgeons standardized form

The ASES was designed as a generic tool to assess patient-related pain and "function/disability" of the entire extremity.[7,12,26] It contains 11 items divided into pain (1 item) and function (10 items). The 10-function items include specific self-care activities (ie, activities of daily living), as well as 1 question on work and 1 question on recreation. The recall period of the items is 1 week. The patient provides a score for each activity for both the right and left shoulder/elbow. The ASES is easy to complete, taking approximately 4 minutes. The pain score is measured on a 10 cm VAS, with 0 = no pain and 10 = pain as bad as it can be. Each function item is scored on a 0 (unable) to 3 (no difficulty) scale. A scoring system for the total score provides weights for pain and activity items, although no rationale has been described for the weighting.[8] Higher scores indicate less pain and disability.[9] There is also a physician-assessment section for ROM, strength, instability, and documentation of specific physical signs, none of which are included in the score.[11,26]

The ASES was developed by a "research" committee by reviewing all published shoulder questionnaires available at the time and including their own ideas.[8] Item selection was not described.[8] Revisions were made based on suggestions of clinicians who were encouraged to use the instrument. Further item reduction techniques and factor analysis is not described.

For internal consistency, Cronbach's alpha ranges from 0.61 to 0.96.[9,19] The range of published ICCs for the ASES is 0.84 to 0.96.[19,27,28] SEM and MDC scores have been reported on a limited basis.[7,9] Construct validity was demonstrated by high ICCs ($r > 0.70$) with CSS, SST, SPADI, and DASH.[7,12,13,19] Correlations were also appropriately weak with dissimilar measures, such as the SF-36.[9] The ASES has been demonstrated to be able to discriminate between patients with "high and low levels of shoulder disability,"[7,9] as well as being responsive to change in patients with orthopedic disorders of the shoulder.[9,26,28,29] Normative data have been published and some cross-cultural evaluation has been performed in German and Italian patients.[17,30,31]

Although the ASES has undergone some limited test-retest reliability and validity testing,[9] it has never been assessed for these properties in patients with head and neck cancer. Although others have reported it to be easy to score,[9] it does contain a VAS, which requires conversion to a numeric score, thus making scoring more time-consuming.[15] Kirkley et al[8] note that the response options from 0 to 3 could potentially limit the sensitivity for responsiveness testing. Last, because the ASES assesses both shoulder and elbow difficulties, patients with elbow problems unrelated to the neck dissection may report difficulties.

Disabilities of the Arm, Shoulder, and Hand questionnaire

The DASH questionnaire is a generic measure of "disability and symptoms" related to any condition of any joint of the upper extremity.[8,32] It is a 30-item questionnaire (21 physical function items, 6 symptom items, and 3 social/role function items) with 2 optional 4-item modules designed to measure the impact of upper extremity "disability" on work (work module) or playing sports or musical instruments (sports and performing arts module). The recall period for items is 1 week. The DASH is easy to use and takes less than 13 minutes to complete.[15] A 5-point scale is used for each item with 1 = no difficulty and 5 = extremely difficult. A disability/symptom score is easily determined, which is converted to a score out of 100, with higher scores representing greater disability. The work and sports/performing arts module are each scored separately.

Item generation was carried out by first reviewing the literature and producing an item pool, which were reviewed by a collaborative group. Items were stripped of attribution to a specific disorder. Items that were repetitive or unrelated to the upper extremity were eliminated. The reduced list was sent to clinician content experts for their input on face and content validity and importance of items. Initial items were formatted into a questionnaire and pilot tested on 20 patients with upper limb problems to ensure readability, absence of ambiguity, and understanding of scale and content. Item reduction was further performed by field-testing in over 400 patients across 20 centers worldwide. Factor analysis was performed demonstrating that the DASH main module could be scored in 1 dimension. A clinimetric reduction also was performed by including formal patient input using an importance and severity questionnaire.[33]

The DASH has undergone extensive testing of its psychometric properties.[7,9,13,15,32–34] Cronbach's alpha for

internal consistency ranges from 0.92 to 0.98.[9,19] Test–retest reliability studies that have been performed in different patient samples have demonstrated ICCs ranging from 0.77 to 0.98.[7,9,13,15] Both the MDC scores and SEMs have been defined for a number of patient groups and interventions. For convergent validity, Pearson's or Spearman's correlation of DASH scores with the ASES, SPADI, and CSS exceeded 0.70.[7,9,34] Divergent validity was also demonstrated with weak correlations with dissimilar measures.[19,34] The DASH has been shown to be able to discriminate between different levels of disease and condition severity (both patient and clinician-rated).[7,33–35] It has also been found to be sensitive enough to detect and differentiate small and large changes of "disability" over time in patients with upper extremity MSK disorders.[34] Although the DASH has not undergone formal assessment of its measurement properties in patients after neck dissection, the DASH has been demonstrated to be able to discriminate between patients with and without neural injury, as well as between those with and without recovery after nerve injury.[36,37] The DASH has also undergone extensive cross-cultural evaluation and normative data has also been determined.[19]

Although the DASH has undergone extensive assessment of its psychometric properties, these properties have not been formally assessed in patients with head and neck cancer. The 2 potential limitations are that it measures a region (arm, shoulder, and hand) rather than being a shoulder-specific questionnaire and that it has more items than most of the other questionnaires and, therefore, takes longer to complete.

Simple Shoulder Test

The SST is a 12-item questionnaire that was developed to assess functional improvement resulting from a specified procedure for a given diagnosis and to characterize the severity of the condition.[8] The items consist of a subjective component, such as questions asking about shoulder comfort (both at rest and sleeping) and questions about performing specific activities. There are also items that require a patient to actually perform physical activities, such as lifting specified weights to specified heights. The response option for each item is a "yes" or "no" response. The recall period is at the actual time of completion.[19] It is easy to complete and can be administered in 3 minutes. A score of 1 is assigned to a "yes" response and score of 0 for a "no" response, with a maximum score of 12. The score is converted to a percentage score out of 100.

Items were generated by reviewing shoulder questionnaires and the most frequent complaints of patients observed in a shoulder practice. There is no further description of how items were selected or reduced and how content validity was determined.[8,19] Factor analysis demonstrated 2 factors despite a 1-factor total score.[19]

For internal consistency, Cronbach's alpha is 0.85.[19] There has been limited assessment of test-retest reliability with reported ICCs of 0.97 and 0.99.[19] There is no reported SEM or MDC score. For convergent validity, correlations (Pearson's or Spearman's) of the SST with the SPADI, ASES, CSS, and DASH were all ≥0.70. On initial assessment, the SST was able to differentiate

between patients with varying shoulder conditions and a sample of patients without any shoulder disorders.[8] There have been a few studies demonstrating responsiveness.[19] Cross-cultural evaluation has been performed in Italian patients who underwent neck dissection.[14]

The major limitations, as pointed out by Kirkley et al and other authors,[7–9,19] is that it is unlikely to be responsive to change because of dichotomous responses, similarly, it is unlikely to have good discrimination ability to differentiate between patients with varying severity of the same condition. There is also no normative data. Some of the activity items ask about perceived ability to perform the task rather than the actual ability to carry out the task. There are a limited number of functional items. There has been no assessment of construct validity in the patients with head and neck cancer undergoing neck dissection. Last, it also incorporates physical evaluation with subjective assessment.

DISCUSSION

When choosing a PRO measure to assess shoulder disability, the measure should be one that has undergone accepted methodology in terms of development and also testing of its psychometric properties, preferably in the population in whom it is intended to be used. The purpose of this review was to highlight the strengths and weaknesses of the PROs that have been used to assess shoulder disability after neck dissection. The PROs chosen to undergo assessment were the ones that have been used in the head and neck cancer literature. These also represent the most common shoulder disability tools used in the literature. A number of other self-report questionnaires have been used in the MSK literature; however, they were developed for specific MSK shoulder disorders or surgeries. These include the Bankart Repair Scoring System, Western Ontario Osteoarthritis Shoulder Index, Western Ontario Rotator Cuff Index, Rotator Cuff Quality of Life Measure, Oxford Shoulder Score, UCLA Shoulder Scale, Oxford Shoulder Instability Questionnaire, and the Western Ontario Shoulder Instability Index. Although many of these instruments have undergone appropriate development and psychometric assessment in their respective populations of interest, the major concern is that the construction of the item pool, item selection, and reduction for these disease-specific questionnaires were performed for the specified population or procedure. The etiology of shoulder impairments and resultant morbidity after neck dissection is different from the MSK etiologies that the above-mentioned instruments were developed to assess. Shoulder impairments after neck dissection result from injury to the motor innervation of the muscles of the shoulder girdle, whereas, in the other cases, impairments may result from injury to the muscles, joint, bones, or tendons of the shoulder girdle. Patient demographics (such as age, sex, comorbidity, socioeconomic status, occupation, commonly performed recreational, and daily activities) and other associated impairments (such as speech and swallowing) may also differ in patients with head and neck cancer compared with patients with MSK diagnoses who were used in the development and assessment of the measurement properties of these instruments.

There are many strengths and weaknesses of each of the shoulder disability questionnaires used in the head and neck cancer literature. The one chosen should depend on the objectives of the study. That being said, acceptance of well-designed patient-reported measure(s) of disability would help facilitate future studies, such as an evaluation of interventions aimed at preventing and rehabilitating shoulder problems after neck dissection. Efforts should be focused on using well-designed questionnaires that have been assessed in this patient population rather than developing or using other questionnaires. Although all the measures used in this review have undergone at least some form of assessment of their psychometric properties, only the NDII was specifically designed and assessed in this patient population. However, demonstration of additional properties of the NDII, such as responsiveness, SEMs, and MDCs, is still required. In comparison with the remaining questionnaires reviewed, the DASH seems to have undergone the most extensive development, as well as assessment of its psychometric properties. However, its psychometric properties have never been formally assessed in patients undergoing neck dissection. Further evaluation in this patient population is required.

REFERENCES

1. Goldstein DP, Ringash J, Bissada E, et al. Scoping review of the literature on shoulder impairments and disability after neck dissection. *Head Neck* 2013 [Epub ahead of print].
2. Gliklich RE, Goldsmith TA, Funk GF. Are head and neck specific quality of life measures necessary? *Head Neck* 1997;19:474–480.
3. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res* 2010;19:539–549.
4. Taylor RJ, Chepeha JC, Teknos TN, et al. Development and validation of the neck dissection impairment index: a quality of life measure. *Arch Otolaryngol Head Neck Surg* 2002;128:44–49.
5. Portney LG, Watkins MP, editors. Statistical measures of reliability. Foundations of Clinical Research, 2nd ed. New Jersey: Prentice Hall Health; 2000. pp 557–586.
6. Roach KE, Budiman–Mak E, Songsiridej N, Lertratanakul Y. Development of a shoulder pain and disability index. *Arthritis Care Res* 1991;4:143–149.
7. Roy JS, MacDermid JC, Woodhouse LJ. Measuring shoulder function: a systematic review of four questionnaires. *Arthritis Rheum* 2009;61:623–632.
8. Kirkley A, Griffin S, Dainty K. Scoring systems for the functional assessment of the shoulder. *Arthroscopy* 2003;19:1109–1120.
9. McClure P, Michener LA. Measures of adult shoulder function. *Arthritis Rheum* 2003;49:S50–S58.
10. Breckenridge JD, McAuley JH. Shoulder Pain and Disability Index (SPADI). *J Physiother* 2011;57:197.
11. Kirkley A, Griffin S, Dainty K. Scoring systems for the functional assessment of the shoulder. *Arthroscopy* 2003;19:1109–1120.
12. Beaton DE, Richards RR. Measuring function of the shoulder. A cross-sectional comparison of five questionnaires. *J Bone Joint Surg Am* 1996;78:882–890.
13. Michener LA, Leggin BG. A review of self-report scales for the assessment of functional limitation and disability of the shoulder. *J Hand Ther* 2001;14:68–76.
14. Marchese C, Cristalli G, Pichi B, et al. Italian cross-cultural adaptation and validation of three different scales for the evaluation of shoulder pain and dysfunction after neck dissection: University of California – Los Angeles (UCLA) Shoulder Scale, Shoulder Pain and Disability Index (SPADI) and Simple Shoulder Test (SST). *Acta Otorhinolaryngol Ital* 2012;32:12–17.
15. Bot SD, Terwee CB, van der Windt DA, Bouter LM, Dekker J, de Vet HC. Clinimetric evaluation of shoulder disability questionnaires: a systematic review of the literature. *Ann Rheum Dis* 2004;63:335–341.
16. Desai AS, Dramis A, Hearnden AJ. Critical appraisal of subjective outcome measures used in the assessment of shoulder disability. *Ann R Coll Surg Engl* 2010;92:9–13.
17. Angst F, Goldhahn J, Pap G, et al. Cross-cultural adaptation, reliability and validity of the German Shoulder Pain and Disability Index (SPADI). *Rheumatology (Oxford)* 2007;46:87–92.
18. Jamnik H, Spevak MK. Shoulder Pain and Disability Index: validation of Slovene version. *Int J Rehabil Res* 2008;31:337–341.
19. Angst F, Schwyzer HK, Aeschlimann A, Simmen BR, Goldhahn J. Measures of adult shoulder function: Disabilities of the Arm, Shoulder, and Hand Questionnaire (DASH) and its short version (QuickDASH), Shoulder Pain and Disability Index (SPADI), American Shoulder and Elbow Surgeons (ASES) Society standardized shoulder assessment form, Constant (Murley) Score (CS), Simple Shoulder Test (SST), Oxford Shoulder Score (OSS), Shoulder Disability Questionnaire (SDQ), and Western Ontario Shoulder Instability Index (WOSI). *Arthritis Care Res (Hoboken)* 2011;63 Suppl 11:S174–S188.
20. van der Heijden GJ, Leffers P, Bouter LM. Shoulder disability questionnaire design and responsiveness of a functional status measure. *J Clin Epidemiol* 2000;53:29–38.
21. de Winter AF, van der Heijden GJ, Scholten RJ, van der Windt DA, Bouter LM. The Shoulder Disability Questionnaire differentiated well between high and low disability levels in patients in primary care, in a cross-sectional study. *J Clin Epidemiol* 2007;60:1156–1163.
22. Conboy VB, Morris RW, Kiss J, Carr AJ. An evaluation of the Constant–Murley shoulder assessment. *J Bone Joint Surg Br* 1996;78:229–232.
23. Chepeha DB, Taylor RJ, Chepeha JC, et al. Functional assessment using Constant's Shoulder Scale after modified radical and selective neck dissection. *Head Neck* 2002;24:432–436.
24. Kirkley A, Alvarez C, Griffin S. The development and evaluation of a disease-specific quality-of-life questionnaire for disorders of the rotator cuff: The Western Ontario Rotator Cuff Index. *Clin J Sport Med* 2003;13: 84–92.
25. Constant CR, Murley AH. A clinical method of functional assessment of the shoulder. *Clin Orthop Relat Res* 1987;214:160–164.
26. Richards RR, An KN, Bigliani LU, et al. A standardized method for the assessment of shoulder function. *J Shoulder Elbow Surg* 1994;3:347–352.
27. Kocher MS, Horan MP, Briggs KK, Richardson TR, O'Holleran J, Hawkins RJ. Reliability, validity, and responsiveness of the American Shoulder and Elbow Surgeons subjective shoulder scale in patients with shoulder instability, rotator cuff disease, and glenohumeral arthritis. *J Bone Joint Surg Am* 2005;87:2006–2011.
28. Michener LA, McClure PW, Sennett BJ. American Shoulder and Elbow Surgeons Standardized Shoulder Assessment Form, patient self-report section: reliability, validity, and responsiveness. *J Shoulder Elbow Surg* 2002; 11:587–594.
29. Angst F, Goldhahn J, Drerup S, Aeschlimann A, Schwyzer HK, Simmen BR. Responsiveness of six outcome assessment instruments in total shoulder arthroplasty. *Arthritis Rheum* 2008;59:391–398.
30. John M, Angst F, Awiszus F, King GJ, MacDermid JC, Simmen BR. The American Shoulder and Elbow Surgeons Elbow Questionnaire: cross-cultural adaptation into German and evaluation of its psychometric properties. *J Hand Ther* 2010;23:301–313; quiz 314.
31. Padua R, Padua L, Ceccarelli E, Bondi R, Alviti F, Castagna A. Italian version of ASES questionnaire for shoulder assessment: cross-cultural adaptation and validation. *Musculoskelet Surg* 2010;94 Suppl 1:S85–S90.
32. Hudak PL, Amadio PC, Bombardier C. Development of an upper extremity outcome measure: the DASH (disabilities of the arm, shoulder and hand) [corrected]. The Upper Extremity Collaborative Group (UECG). *Am J Ind Med* 1996;29:602–608.
33. Solway S, Beaton DE, McConnell S, Bombardier C. The DASH outcome measure user's manual (2nd edition). Toronto, Ontario, Canada: Institute for Work and Health; 2002.
34. Beaton DE, Katz JN, Fossel AH, Wright JG, Tarasuk V, Bombardier C. Measuring the whole or the parts? Validity, reliability, and responsiveness of the Disabilities of the Arm, Shoulder and Hand outcome measure in different regions of the upper extremity. *J Hand Ther* 2001;14:128–146.
35. Getahun TY, MacDermid JC, Patterson SD. Concurrent validity of patient rating scales in assessment of outcome after rotator cuff repair. *J Musculoskelet Res* 2000;4:119–127.
36. Bushnell BD, McWilliams AD, Whitener GB, Messer TM. Early clinical experience with collagen nerve tubes in digital nerve repair. *J Hand Surg Am* 2008;33:1081–1087.
37. Novak CB, Anastakis DJ, Beaton DE, Katz J. Patient-reported outcome after peripheral nerve injury. *J Hand Surg Am* 2009;34:281–287.