

Assessment of the Disabilities of the Arm, Shoulder, and Hand (DASH) questionnaire for use in patients after neck dissection for head and neck cancer

David P. Goldstein, MD,^{1*} Jolie Ringash, MD,² Jonathan C. Irish, MD,¹ Ralph Gilbert, MD,¹ Patrick Gullane, MD,¹ Dale Brown, MD,¹ Wei Xu, PhD,³ Ryan Del Bel,³ Douglas Chepeha, MD,⁴ Aileen M. Davis, PhD⁵

¹Department of Otolaryngology–Head and Neck Surgery, University Health Network, Princess Margaret Cancer Center, University of Toronto, Toronto, Ontario, Canada, ²Department of Radiation Oncology, University Health Network, Princess Margaret Cancer Center, University of Toronto, Toronto, Ontario, Canada, ³Department of Biostatistics, Princess Margaret Cancer Center, University of Toronto, Toronto, Ontario, Canada, ⁴Department of Otolaryngology–Head and Neck Surgery, University of Michigan, Ann Arbor Michigan, ⁵Health Care and Outcomes Research, Toronto Western Research Institute, University Health Network, University of Toronto, Toronto, Ontario, Canada.

Accepted 20 December 2013

Published online 26 September 2014 in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/hed.23593

ABSTRACT: *Background.* In this cross-sectional study, the sensibility, test-retest reliability, and validity of the Disabilities of the Arm, Shoulder, and Hand (DASH) questionnaire were assessed in patients who underwent neck dissection.

Methods. Sensibility was assessed with a questionnaire. Test-retest reliability was performed with completion of the DASH questionnaire 2 weeks after initial completion; validity, by evaluating differences in scores between patients undergoing different types of neck dissections and correlating DASH scores with Neck Dissection Impairment Index (NDII) scores.

Results. The DASH questionnaire met sensibility criteria. For test-retest reliability analysis, the intraclass coefficient was 0.91. The DASH questionnaire showed differences between patients who underwent acces-

sory nerve-sacrifice and nerve-sparing neck dissection. DASH questionnaire scores strongly correlated with NDII scores ($r = -0.86$).

Conclusion. Although this study provides preliminary data on some psychometric properties of the DASH questionnaire in patients who have undergone a neck dissection, further assessment of responsiveness and other properties are required. © 2014 Wiley Periodicals, Inc. *Head Neck* 37: 234–242, 2015

KEY WORDS: shoulder impairments, shoulder disability, shoulder scales, Disabilities of the Arm, Shoulder, and Hand (DASH) questionnaire, Neck Dissection Impairment Index (NDII), Shoulder Pain and Disability Index

INTRODUCTION

Several different patient-based questionnaires have been used to assess shoulder disability after neck dissection in patients with head and neck cancer. Many of these would not meet contemporary standards for measurement^{1,2} and/or were not specifically developed or evaluated for use in the head and neck cancer population.³ The Neck Dissection Impairment Index (NDII) is the only questionnaire developed in this specific patient population,⁴ however, it has not been uniformly used as the primary outcome measure for shoulder disability.

After an extensive literature search and critical review of existing shoulder disability measures, the Disabilities of the Arm, Shoulder and Hand (DASH) questionnaire was identified as a potential patient-reported outcome measure to study for use in patients who have undergone neck dissection. The DASH questionnaire has undergone

extensive test-retest reliability, validity, and responsiveness testing.^{5–7} However, the robustness of this measure in terms of these measurement properties has not been demonstrated in the patients with head and neck cancer. Patients with head and neck cancer have unique problems that require investigations specific to their disease-related issues. The etiology of the activity limitations of the shoulder in patients with head and neck cancer, such as accessory nerve injury, differs from that seen in the samples used for the development of most other instruments to assess shoulder impairments (eg, rotator cuff injury and shoulder instability). Accessory nerve injury is relatively unique to head and neck surgery and deliberate sacrifice of the nerve is relatively specific to patients requiring neck dissection for head and neck cancer. The primary purpose of this study was to assess the sensibility, test-retest reliability, and validity of the DASH questionnaire in patients who have undergone a neck dissection for head and neck cancer.

PATIENTS AND METHODS

A cross-sectional study was undertaken after institutional ethics review board approval was obtained. The sample included patients who underwent radical neck dissection, accessory nerve sparing modified radical neck dissection, or selective neck dissection at the Princess

*Corresponding author: D. P. Goldstein, Department of Otolaryngology–Head and Neck Surgery, University Health Network, Princess Margaret Cancer Center, Princess Margaret Hospital, Room 3-952, 610 University Avenue, Toronto Ontario, M5G 2M9, Canada. E-mail: david.goldstein@uhn.on.ca

Contract grant sponsor: This work was supported by grants from Cancer Care Ontario, Ontario Cancer Research Network, and Canadian Institutes for Health Research.

Margaret Cancer Center, University of Toronto, between January 1998 and July 2005 for primary (nonmetastatic) head and neck cancer of the upper aerodigestive tract (oral cavity, oropharynx, nasopharynx, or larynx), skin, thyroid, or salivary glands.

Participants were >19 years of age and had a neck dissection either as part of their primary treatment or for persistent or recurrent neck disease after radiotherapy (with or without chemotherapy). All patients included were disease-free at the time of inclusion in the study. Similar to the inclusion criteria used for the development and assessment of the NDII,⁴ all participants were 11 or more months postsurgery. Because much of the clinical concern regarding morbidity after neck dissection surrounds long-term shoulder disability, patients who were more remote from their surgery were deliberately accrued.

Patients with the following criteria were excluded: (1) presence of distant metastases or incurable disease; (2) salvage neck dissection for recurrent neck disease after previous neck dissection; (3) neck dissection with resection of the splenius capitus, trapezius, or levator scapulae muscles; (4) ipsilateral scapula osteocutaneous free flap reconstruction; (5) history of any shoulder surgery or trauma, rheumatologic disorders requiring disease modifying agents, or prior neurological or cerebrovascular disease affecting the upper extremity (6) bilateral neck dissection; and (7) inability to read or write English.

A list of patients who underwent neck dissection was generated from a prospective head and neck surgical registry. All medical charts of patients who underwent radical neck dissection ($n = 255$) were reviewed for eligibility because of the small number of radical neck dissections performed per year and the high mortality rate of these patients. Given the large number of patients who underwent modified radical neck dissection or selective neck dissection during the time period ($n = 706$), a sample of these patients was obtained by reviewing 300 consecutive medical charts starting from patients treated in January 2005 and sequentially working backward in time in order to determine eligibility. Of these patients, 141 (84 selective neck dissection and 57 modified radical neck dissection) were eligible for inclusion in this study. Using a random number generator, 90 of these patients (45 who had selective neck dissection and 45 who had modified radical neck dissection) were randomly chosen to be eligible for inclusion. For sensibility testing, physicians were eligible for inclusion if they had a surgical practice specializing in the management of head and neck cancer. A list of surgeons was generated from the faculty at academic centers in both Canada and the United States that offer fellowships in head and neck oncologic surgery.⁸

Eligible patient participants were mailed a questionnaire package containing the sensibility questionnaire, the DASH questionnaire, and the NDII. Physician participants were mailed a package with a clinical practice survey, the DASH questionnaire, and sensibility questionnaire. Two weeks after the initial mailing, a postcard reminder was mailed, followed by mailing a second questionnaire package 2 weeks later to those participants from whom a

returned package had not been received (ie, a modified Dillman approach).^{9,10} Data collected for patient participants were from a combination of medical chart review and mailed questionnaire packages. Chart review of non-responders was not performed because failure to return a package was considered by the institutional ethics board as the patient's refusal to participate in the study.

Sensibility, which consists of an aggregate of properties that make up the common sense aspect of an instrument (ie, face and content validity, comprehensibility, replicability, suitability of scale, and ease of usage)¹¹ was assessed by surveying the patients with head and neck cancer and their surgeons. For assessment of test-retest reliability, a package containing the DASH questionnaire was mailed out approximately 2 weeks after the first package was returned. A change in status form was included in order to determine if the participant's shoulder disability or function was the same, better, or worse since the time they completed the first questionnaire. Participants who reported a change in status were excluded from the test-retest reliability analysis. Known-group validity was assessed by evaluating differences in the DASH questionnaire scores between patients who underwent different types of neck dissections. Convergent validity was evaluated by correlating DASH questionnaire scores with NDII scores.

Measures

Sensibility questionnaire. The sensibility questionnaire was developed based on a sensibility questionnaire designed by Rowe and Oxman.¹² It contained 9 questions that assessed clarity of instructions and questions, ease of usage, whether the content was appropriate, any missing items or issues, and if the DASH questionnaire met its overall goal of assessing shoulder disability after neck dissection. A 7-point response key was provided for each question, with 1 representing the most negative response and 7 being the most positive response. Items were scored individually and space was provided after each question for additional comments.

The DASH questionnaire. The DASH questionnaire (Figure 1) was designed to measure physical symptoms, activity limitations, and participation restrictions related to any condition of any joint of the upper extremity.¹³ 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). Each item is rated with a score from 1 (least disability) to 5 (most disability). For the disability/symptom module of the DASH questionnaire, all 30-response scores are added together, producing a raw score, which is then transformed into a score of 100 maximum. The sports/performing arts and the work module are each scored separately. For all the modules, a higher score indicates greater disability. Missing data were handled according to criteria outlined in the user manual. Ninety percent of the items needed to be completed in order to include a respondents DASH questionnaire scores.¹⁴

DISABILITIES OF THE ARM, SHOULDER AND HAND

Please rate your ability to do the following activities in the last week by circling the number below the appropriate response.

	NO DIFFICULTY	MILD DIFFICULTY	MODERATE DIFFICULTY	SEVERE DIFFICULTY	UNABLE
1. Open a tight or new jar.	1	2	3	4	5
2. Write.	1	2	3	4	5
3. Turn a key.	1	2	3	4	5
4. Prepare a meal.	1	2	3	4	5
5. Push open a heavy door.	1	2	3	4	5
6. Place an object on a shelf above your head.	1	2	3	4	5
7. Do heavy household chores (e.g., wash walls, wash floors).	1	2	3	4	5
8. Garden or do yard work.	1	2	3	4	5
9. Make a bed.	1	2	3	4	5
10. Carry a shopping bag or briefcase.	1	2	3	4	5
11. Carry a heavy object (over 10 lbs).	1	2	3	4	5
12. Change a lightbulb overhead.	1	2	3	4	5
13. Wash or blow dry your hair.	1	2	3	4	5
14. Wash your back.	1	2	3	4	5
15. Put on a pullover sweater.	1	2	3	4	5
16. Use a knife to cut food.	1	2	3	4	5
17. Recreational activities which require little effort (e.g., cardplaying, knitting, etc.).	1	2	3	4	5
18. Recreational activities in which you take some force or impact through your arm, shoulder or hand (e.g., golf, hammering, tennis, etc.).	1	2	3	4	5
19. Recreational activities in which you move your arm freely (e.g., playing frisbee, badminton, etc.).	1	2	3	4	5
20. Manage transportation needs (getting from one place to another).	1	2	3	4	5
21. Sexual activities.	1	2	3	4	5

FIGURE 1. Items in the main module of the Disabilities of the Arm, Shoulder, and Hand Questionnaire (DASH) Questionnaire. Republished with permission of Institute for Work & Health, copyright.

The Neck Dissection Impairment Index. The NDII is a 10-item self-administered questionnaire, which was designed to assess “quality of life related to shoulder dysfunction” after neck dissection.⁴ For the NDII, a 5-point response option (1 to 5) is provided for each item, with 1 being the most disability and 5 the least disability. The

responses for all 10 items are added together to produce a raw score, which is then transformed to a score of 100 maximum. Higher scores represent less disability. Techniques for handling missing data are not described for the NDII and we therefore used the same rules as for the DASH questionnaire.

DISABILITIES OF THE ARM, SHOULDER AND HAND

sample only

	NOT AT ALL	SLIGHTLY	MODERATELY	QUITE A BIT	EXTREMELY
22. During the past week, to what extent has your arm, shoulder or hand problem interfered with your normal social activities with family, friends, neighbours or groups? (circle number)	1	2	3	4	5

	NOT LIMITED AT ALL	SLIGHTLY LIMITED	MODERATELY LIMITED	VERY LIMITED	UNABLE
23. During the past week, were you limited in your work or other regular daily activities as a result of your arm, shoulder or hand problem? (circle number)	1	2	3	4	5

Please rate the severity of the following symptoms in the last week. (circle number)

	NONE	MILD	MODERATE	SEVERE	EXTREME
24. Arm, shoulder or hand pain.	1	2	3	4	5
25. Arm, shoulder or hand pain when you performed any specific activity.	1	2	3	4	5
26. Tingling (pins and needles) in your arm, shoulder or hand.	1	2	3	4	5
27. Weakness in your arm, shoulder or hand.	1	2	3	4	5
28. Stiffness in your arm, shoulder or hand.	1	2	3	4	5

	NO DIFFICULTY	MILD DIFFICULTY	MODERATE DIFFICULTY	SEVERE DIFFICULTY	SO MUCH DIFFICULTY THAT I CAN'T SLEEP
29. During the past week, how much difficulty have you had sleeping because of the pain in your arm, shoulder or hand? (circle number)	1	2	3	4	5

	STRONGLY DISAGREE	DISAGREE	NEITHER AGREE NOR DISAGREE	AGREE	STRONGLY AGREE
30. I feel less capable, less confident or less useful because of my arm, shoulder or hand problem. (circle number)	1	2	3	4	5

DASH DISABILITY/SYMPTOM SCORE = $\frac{(\text{sum of } n \text{ responses})}{n} - 1 \times 25$, where n is equal to the number of completed responses.

A DASH score may not be calculated if there are greater than 3 missing items.



© INSTITUTE FOR WORK & HEALTH 2006. ALL RIGHTS RESERVED

FIGURE 1. (Continued)

Sample size calculations

Because the DASH questionnaire was not used previously in the patients with head and neck cancer, the mag-

nitude of score differences between groups was not known, and, therefore, sample size calculations were not performed for sensibility or validity testing. The sample

size was estimated at 30 patients in each neck dissection group, which was considered a reasonable sample size to show the pattern of expected differences. For the test-retest reliability analysis, 47 analyzable cases were required based on achieving an intraclass correlation coefficient (ICC) of 0.80 with a lower bound 95% confidence interval (CI) of 0.70.¹⁵ Sample size for the clinician component of sensibility analysis was conducted based on obtaining a sample of at least 50 experts.

Statistical analyses

All analyses were completed using the SPSS v14 (Statistical Package for the Social Sciences for Windows, SPSS, Chicago, IL). Sociodemographic, clinicopathologic, and treatment-related data were summarized using descriptive summary statistics as appropriate for the type of data. Differences between the neck dissection groups in these variables were assessed using chi-square analysis for categorical data and 1-way analysis of variance (ANOVA) or nonparametric tests for interval data.

For sensibility analysis a priori criteria, the Rowe and Oxman¹² criteria were used to assess the sensibility of the DASH questionnaire. The DASH questionnaire was judged to be sensible if: (1) the mean scores provided by the patients were ≥ 5 for at least 7 of 9 ($>80\%$) of the items in the sensibility questionnaire; (2) mean scores for the physicians were ≥ 4 for at least 7 of 9 of the items on the sensibility questionnaire; and (3) <3 of the items were given a mean score of ≤ 3 by either group.¹² Content analysis of the written responses for each question was performed by identifying similar themes for patients and physicians.

Test-retest reliability for the DASH questionnaire was calculated using the ICC (model 2) with 95% CIs.^{16,17} Weighted kappa scores could not be used to assess reliability coefficients for individual items in the DASH questionnaire because the assumption for weighted kappa scores was violated. Individual item test-retest reliability was therefore assessed by determining for each item the number of patients whose numerical response/score on the 5-point response option was identical between the test and retest response (percent agreement); or differed by 1 or more (in either direction). As some variability is expected between tests, a 1 score difference was still considered to be an indication of test-retest reliability for individual questions.

Differences in DASH questionnaire scores between selective neck dissection, modified radical neck dissection, and radical neck dissection groups were determined by using a 1-way ANOVA Kruskal-Wallis nonparametric ANOVA and a post-hoc Mann-Whitney *U* test. Multi-variable analysis (MVA) of variables (type of neck dissection, time from surgery, radiation, and age) associated with DASH questionnaire scores was performed using linear regression analysis. The covariates chosen to be entered into the MVA were variables that were either used in the validation of the NDII⁴ or felt to be potential confounders of shoulder disability. Body mass index, which was assessed with the NDII, was not included in the MVA, as this data was not recorded. Association between the DASH questionnaire and NDII scores was

determined using the Spearman's rank correlation coefficient.

RESULTS

One hundred fifty-three patients were mailed a package (45 selective neck dissections, 45 modified radical neck dissections, and 63 radical neck dissections). Six packages were returned to sender and 10 patients had died. One hundred nine of the 137 patients who were presumed to receive a package returned a completed package (response rate of 79.6%). Thirteen of these patients reported having bilateral neck dissections and were excluded from analysis. Therefore, a total of 96 patients who returned a package were eligible for inclusion. Sociodemographic, clinicopathologic, and treatment-related data for the entire cohort and each neck dissection group are presented in Table 1. Fifty of the 94 physicians who were sent a package responded (response rate of 53%). Characteristics of physician respondents are presented in Table 2.

Sensibility assessment

Ninety-six patients and 50 physicians were included in the sensibility assessment. Based on the pre-established criteria, the DASH questionnaire was deemed to be sensible by both the patient and the physician respondents. The mean and median scores for all the questions on the sensibility questionnaire were >4 (Table 3). On content analysis, overall, patients and physicians stated that the DASH questionnaire was a comprehensive questionnaire that was easy to use and understand. Some patients and physicians noted that there was no method for handling items deemed not applicable. A few patients felt the questionnaire missed important questions about problems patients face with day-to-day care; examples included "putting on a coat," "dental hygiene," or "shaving," or problems at work, such the ability to use a computer. These latter patients reported lower accuracy and slower performance on a computer after their surgery. One area of self-care specific to patients with head and neck cancer that physicians felt should be included in the questionnaire was "difficulty with tracheotomy or gastrostomy tube care."

Test-retest reliability

Forty-four of 52 patients who were mailed a package for test-retest reliability returned a completed package (response rate of 84.6%). Thirty-five patients were eligible for inclusion (9 were excluded because of too many missing items ($n = 2$), reporting having had bilateral neck dissection ($n = 5$), or a change in status ($n = 2$). Both patients who reported a change in status were patients who had a modified radical neck dissection >24 months prior and reported improvement between the test and retest. The mean (median) time between receiving the patient's initial response and the re-test mailing was 2.8 (2.5) weeks. Internal consistency as measured by Cronbach's alpha on the entire sample (ie, not the test-retest reliability sample) was 0.959. The mean (median, SD) DASH questionnaire score for the initial test sample was 12.9 (range, 7.5–17.7) and for the retest sample was 14.2

TABLE 1. Demographic, clinicopathologic, and treatment-related data for the entire cohort and neck dissection group.

Variable	Total group (n = 96)	Selective neck dissection (n = 34)	Modified radical neck dissection (n = 31)	Radical neck dissection (n = 31)	p value
Sex					
Male	61 (63.5%)	17 (50%)	21 (67.7%)	22 (71%)	.167
Female	35 (36.5%)	17 (50%)	10 (32.2%)	9 (29%)	
Age, y					
Mean (median)	62.7 (63)	61.6 (60)	60.9 (64)	65.8 (64)	.302
Tumor site					
SCC UADT	54 (56%)	21 (62%)	10 (32%)	23 (74%)	.001
Skin	18 (18.7%)	4 (12%)	8 (26%)	6 (19%)	
Thyroid	12 (12.5%)	2 (6%)	9 (29%)	1 (3.2%)	
Salivary gland	12 (12.5%)	7 (21%)	4 (13%)	1 (3.2%)	
Adjuvant treatment					
None	37 (38.5%)	17 (50%)	17 (54.8%)	3 (9.7%)	.006
Preoperative RT	21 (21.9%)	4 (12%)	5 (16.1%)	13 (41.9%)	
Preoperative CRT	10 (10.4%)	2 (5.9%)	2 (6.5%)	6 (19.4%)	
Postoperative RT	25 (26%)	10 (29.4%)	6 (19.4%)	8 (25.8%)	
Missing	3 (3%)	1 (2.9%)	1 (3.2%)	1 (3.2%)	
Time from surgery					
Mean, mo	38	29.2	27.8	59.9	.001
Surgery on side of dominant hand					
Yes	43 (44.8%)	15 (44.1%)	12 (38.7%)	16 (51.6%)	.591
No	53 (55.2%)	19 (55.9%)	19 (61.3%)	15 (48.4%)	
Working status					
Unemployed	41 (43%)	15 (44%)	12 (39%)	15 (48.4%)	.943
Employed	48 (50%)	16 (47%)	15 (48%)	16 (51.6%)	
Missing	7 (7%)	3 (9%)	4 (13%)		

Abbreviations: SCC, squamous cell carcinoma; UADT, upper aerodigestive tract (includes oral cavity, oropharynx, hypopharynx, larynx, nasopharynx, unknown primary carcinoma); RT, radiotherapy; CRT, chemoradiotherapy.

(range, 10–16.8). For test-retest reliability, the ICC for the DASH questionnaire main module was 0.91 (95% CI, 0.90–0.98). For individual item test-retest reliability of each of the 30 items on the DASH questionnaire, all 30 items had >86% of respondents providing a test and retest score that was either identical or within 1 score difference.

TABLE 2. Characteristics of physician respondents.

Characteristic	No. of surgeons (%)
Country of practice	
Canada	26 (52)
United States	24 (48)
Head and neck fellowship trained	
No	12 (24)
Yes	37 (74)
Years in practice	
<10	19 (38)
10–20	18 (36)
>20	11 (22)
Percentage of practice devoted to head and neck oncology	
0% to 25%	3 (6)
25% to 50%	2 (4)
50% to 75%	8 (16)
>75%	35 (70)
Total no. of neck dissections performed per year	
<25	3 (6)
26–50	20 (40)
51–75	11 (22)
76–100	6 (12)
>100	6 (12)

Validity analysis

Ninety-one patients were included in the validity analysis. Five were excluded because of too many missing items. The mean (median, SD) DASH questionnaire score for the entire cohort was 17.10 (range, 10.8–18.5) with a range of scores between 0 and 81. For known-group validity, the mean DASH questionnaire main module scores and 95% CI for the selective neck dissection, modified radical neck dissection, and radical neck dissection patients were: 12.05 (95% CI, 6.89–17.2); 13.82 (95% CI, 7.7–19.94), and 26.06 (95% CI, 18.49–33.63), respectively. Boxplots for the scores are presented in Figure 2. There was a statistically significant difference in DASH questionnaire scores between the groups (chi-square = 12.432; df = 2; p value = .002). Using the Mann Whitney U test, the following results between neck dissection groups was found: selective neck dissection versus modified radical neck dissection, p = .54; selective neck dissection versus radical neck dissection, p = .001; and modified radical neck dissection versus radical neck dissection, p = .005. Multivariate analysis on neck dissection was applied adjusting for age, time from surgery, and radiation using linear regression. Type of neck dissection was the only factor associated with DASH questionnaire scores (p = .0057). Age (p = .49), time from surgery (p = .41), and history of radiation (p = .88) were not associated with DASH questionnaire scores. Although adjusting for age and radiation exposure, the only significant difference in DASH questionnaire scores was between radical neck dissection and modified radical neck dissection (p = .03) and radical neck dissection and selective

TABLE 3. Mean and median patient scores for each question on sensibility assessment.

Question	All patients	Physicians
	Mean score; median (SD)	Mean score; median (SD)
1. How easy are the questions to understand?	6.26; 7.00 (1.07)	5.98; 6 (0.85)
2. Were the instructions easy to understand?	6.22; 7 (1.06)	5.92; 6 (0.85)
3. How was length of time to complete?	6.03; 7 (1.22)	4.51; 5 (1.42)
4. Did it include important questions on self-care and daily chores?	5.77; 6 (1.33)	5.77; 6 (0.88)
5. Did it include important questions on effect on work?	5.74; 6 (3.07)	5.52; 5.5 (0.91)
6. Did it include important questions on effect on recreation?	5.93; 6 (1.29)	5.84; 6 (0.77)
7. Do you think the goal of the DASH has been achieved?	5.70; 6 (1.28)	5.76; 6 (1.07)
8. Does it miss important issues about shoulder problems?	6.15; 7 (1.18)	6.02; 6 (1.15)
9. Are the questions too probing (that is, too personal)?	6.75; 7 (0.68)	6.18; 7 (0.86)

Abbreviations: DASH, Disabilities of the Arm, Shoulder, and Hand Questionnaire.

neck dissection ($p = .02$). There was no statistically significant difference in DASH questionnaire scores between selective neck dissection and modified radical neck dissection ($p = .66$) when adjusting for these variables.

The mean (median, SD) for NDII scores was 74.3 (80; 25.79) with a range from 2.5–100. For convergent validity, Spearman's rank correlation coefficient between DASH questionnaire main module and NDII score was $p = -0.86$ ($p < .001$). The inverse correlation relates to the fact that, for this study, the scoring for the 2 instruments was opposite with higher scores on the NDII and lower scores on the DASH questionnaire representing less shoulder disability.

DISCUSSION

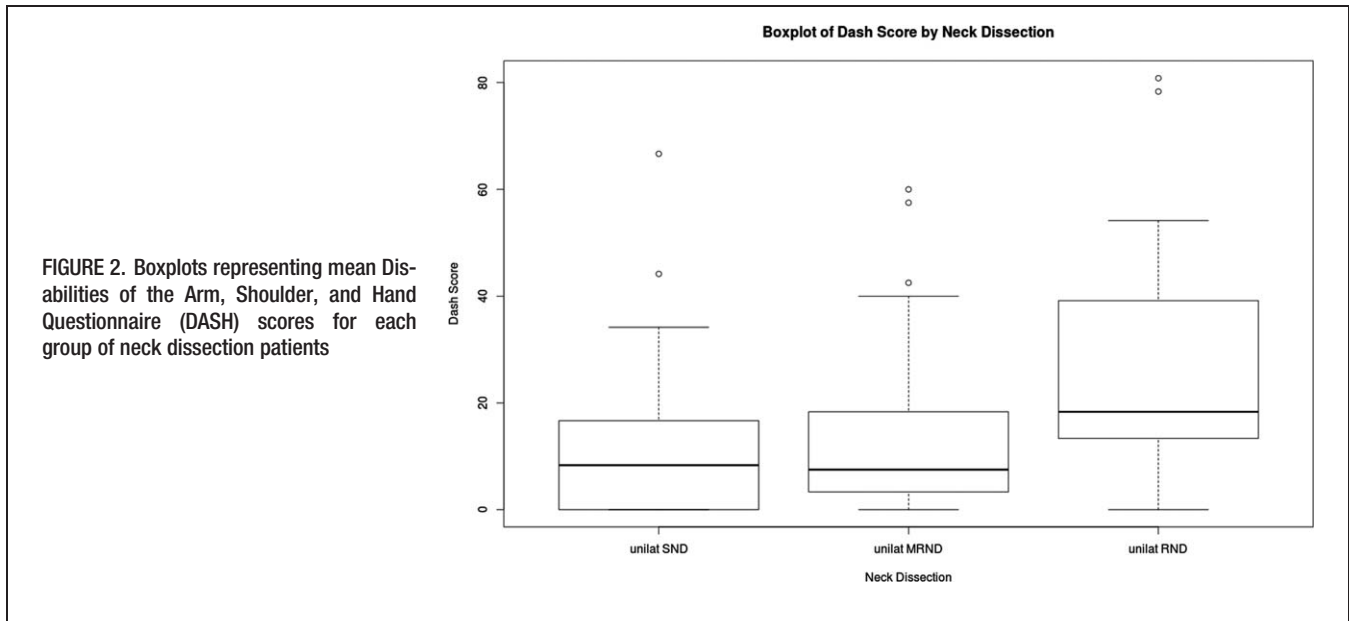
The DASH questionnaire has been one of the most extensively developed and evaluated upper extremity disability questionnaire.^{5,18–20} The DASH questionnaire has shown to be able to discriminate between patients with and without neural injury, as well as between those with and without recovery after nerve injury.^{21,22} It has undergone extensive cross-cultural adaptation. This study provides preliminary support that the DASH questionnaire demonstrated sensibility, test-retest reliability, convergent validity, and known-group validity in patients with head and neck cancer who have undergone neck dissection.

Although the DASH questionnaire was deemed to be a sensible measure of shoulder disability by both patients and physicians, the latter group was more critical. The major concern among the surgeons was that the questionnaire was too long. Patients on the other hand did not feel that questionnaire length was a concern because it can be completed in <10 minutes.²⁰ Some physicians suggested collapsing items into general categories (ie, recreational activity items or items related to activities of daily living), which is similar to the design of the NDII. The risk with this format is that respondents who have difficulty with some but not all activities may feel uncertain about selecting the appropriate answer or may focus only on the activities provided in the examples and may not be able to recall or focus on other commonly per-

formed activities that can be affected by shoulder impairments. Another potential weakness of the DASH questionnaire highlighted by patients and physicians was the lack of a “not applicable” response option. Only a few shoulder disability questionnaires, such as the Shoulder Disability Questionnaire and the American Shoulder and Elbow Scale, have provision for patients to indicate that an item has been left blank because it is “not applicable.”^{23,24} However, “not applicable” responses complicate scoring and interpretation of scores. Last, physicians and patients indicated that there were missing items that should have been included in the DASH questionnaire. Many of the missing activities mentioned were originally included with the initial item generation for the DASH questionnaire but later removed based on psychometric and clinimetric item reduction.¹⁴ In terms of tracheotomy and gastrostomy tube care, the vast majority of patients who undergo a neck dissection do not require either of these procedures, and those who do require a tracheotomy are frequently decannulated before hospital discharge.

The DASH questionnaire main module exhibited excellent test-retest reliability, with an ICC comparable to those reported in other patient samples with different upper extremity conditions, which range from 0.86 to 0.98.^{5,20,25} Interpretation of test-retest reliability depends upon how a measure is to be used or applied. If the DASH questionnaire is going to be used at a group level for research purposes, as was originally intended, lower reliability limits can be tolerated (ie, <0.90).²⁶ If the DASH questionnaire is to be used as a clinical tool upon which individual treatment decisions are to be based, higher test-retest reliability coefficients are required (upwards of 0.90) to ensure limited measurement error and valid interpretation of findings.²⁷ Based on the excellent reliability coefficient for the main module in the current study, the DASH questionnaire meets both requirements, however, further studies are warranted to determine if the DASH questionnaire can or should be used in the clinical setting in patients undergoing neck dissection.

Convergent and divergent validity of the DASH questionnaire have been previously demonstrated in many different patient samples with varying disorders of the upper



extremity.^{5,14,20,25} In these samples, the DASH questionnaire was strongly correlated to other measures of shoulder disability including the Shoulder Pain and Disability Index ($r = 0.85$), Short Form Health Survey (SF-36; physical function and pain domains; $r = 0.73$), Constant's Shoulder Score ($r = -0.76$), American Shoulder and Elbow Surgeons scale ($r = -0.81$), and the visual analogue scale pain intensity ($r = 0.73$).^{5,14,20,25,28} Given that the DASH questionnaire has correlated well with these other measures in many different patient samples, we wanted to determine if the DASH questionnaire would correlate well with a measure of shoulder disability after neck dissection specifically developed for use in the patients with head and neck cancer population (ie, the NDII). Construct validity for the NDII has previously been assessed by correlating scores with Constant's score and the SF-36.⁴ The NDII correlated well with the Constant Shoulder Scale ($r = 0.85$; $p < .001$) and with the SF-36 physical functioning ($r = 0.50$; $p < .001$) and role-physical functioning ($r = 0.60$; $p < .001$) domains.⁴ In the current study, a strong correlation ($r = -0.86$) was found between the scores on the DASH questionnaire and the NDII, implying that the DASH questionnaire and the NDII measure a similar construct.

On known-group validity, the DASH questionnaire was able to discriminate between patients who underwent accessory nerve-sparing neck dissection and those who underwent nerve-sacrificing neck dissection. However, no difference was found in the mean scores between patients who underwent modified radical neck dissection and those who underwent selective neck dissection. It is possible that a difference did not exist between the 2 groups. Other authors have similarly shown that, with time (>6 months), the majority of patients with any nerve-sparing neck dissection experience recovery and a reduction in disability.³ All patients evaluated in the current study were >11 months from surgery and, thus, a significant amount of recovery was expected. To ensure an ability to

detect differences in the DASH questionnaire scores between selective neck dissection and modified radical neck dissection, evaluation could have been performed in the early postoperative period (<6 months) when a greater difference in shoulder morbidity would be expected.³ Confounding variables, such as preexisting shoulder dysfunction, postoperative physiotherapy, and inadvertent accessory nerve injury that were not controlled for also may have been present. Through medical chart review, patients with a documented history of a pre-existing shoulder disorder were excluded. However, this review was limited by the information within the medical charts. Postoperative physiotherapy has been shown to reduce the amount of shoulder impairment after radical neck dissection and to reduce the risk of adhesive capsulitis in patients with accessory nerve neuropraxia.^{29,30} It was not possible to determine the extent of physiotherapy or shoulder exercises from the medical chart review. Although patients may be offered physiotherapy as an out-patient treatment, the clinical impression is that many of our patients do not undergo physiotherapy as it is not covered under the universal health care program. Injury to the accessory nerve at the time of surgery or denervation of the levator scapulae muscle could potentially result in additional shoulder impairments, the extent of which may be influenced by the extent of nerve manipulation, direct injury, or thermal injury from cautery. Similarly, this information was not obtainable from medical chart review.

In addition to the retrospective collection of clinical data and potential confounders, other potential study limitations exist. The physician response rate was only 50%, which is consistent with that reported in the literature.³¹ The sample of head and neck surgeons who were asked to participate in the current study represent a homogenous group of academic surgeons, whose clinical practice is primarily focused on the management of patients with head and neck cancer. Given the homogeneity of the

group, non-response bias is probably unlikely.³² The sample size for test-retest reliability was lower than our estimated sample size. Despite the lower sample size, the ICC CIs for the DASH questionnaire main module were relatively narrow. Last, the psychometric properties of the DASH questionnaire were assessed in patients undergoing neck dissection who were >11 months from surgery. Therefore, the results may not be generalizable to other procedures or may not be applicable to patients assessed at an earlier time point from neck dissection.

CONCLUSION

Although there is a significant amount of literature on shoulder outcomes after head and neck surgery, there is no standard questionnaire used to assess outcome, which makes interpretation of the literature difficult. In addition, further research is needed to help surgeons predict which patients will have shoulder disability after neck dissection and to evaluate if early intervention can help reduce the extent of shoulder-related morbidity. The first step in getting physicians to use 1 or more standardized patient-reported outcome measures would be to ensure that there is adequate literature supporting the use of the instrument(s) for a given objective. The current research provides preliminary support that the DASH questionnaire is a sensible, reliable, and valid measure of shoulder impairment, activity limitations, and participation restrictions in patients with head and neck cancer who have undergone neck dissection. Establishment of other psychometric properties of the DASH questionnaire, such as responsiveness or predictive validity, in the patients with head and neck cancer population, is still required. Additionally, future work could determine if the DASH questionnaire has any advantages over the NDII.

REFERENCES

1. Streiner DL, Norman GR. Health measurement scales: a practical guide to their development and use. 4th ed. New York: Oxford University Press; 2008.
2. Guyatt GH, Kirshner B, Jaeschke R. Measuring health status: what are the necessary measurement properties? *J Clin Epidemiol* 1992;45:1341–1345.
3. 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].
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. 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.
6. 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.
7. Kirkley A, Griffin S, Dainty K. Scoring systems for the functional assessment of the shoulder. *Arthroscopy* 2003;19:1109–1120.
8. American Head and Neck Society. 7th International Conference on Head and Neck Cancer; 2008.
9. Dillman DA. Why choice of survey mode makes a difference. *Public Health Rep* 2006;121:11–13.
10. Dillman DA, Smyth JD, Melani Christian L. Internet, mail, and mixed mode surveys: the tailored design method. Hoboken, NJ: John Wiley & Sons; 2009.
11. Feinstein AR, editor. The theory and evaluation of sensibility. Clinometrics. Westford, MA: Murray Printing Company; 1987. pp 141–166.
12. Rowe BH, Oxman AD. An assessment of the sensibility of a quality-of-life instrument. *Am J Emerg Med* 1993;11:374–380.
13. 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.
14. Solway S, Beaton DE, McConnell S, Bombardier C. The DASH outcome measure user's manual. Toronto, Ontario: Institute for Work and Health; 2002.
15. Kraemer HC, Korner AF. Statistical alternatives in assessing reliability, consistency, and individual differences for quantitative measures: application to behavioral measures of neonates. *Psychol Bull* 1976;83:914–921.
16. Portney LG, Watkins MP, editors. Foundations of clinical research: applications to practice. 2nd ed. New York: Prentice Hall; 2000. pp 557–586.
17. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979;86:420–428.
18. 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.
19. Kirkley AA, Griffin S, Dainty K. Scoring systems for the functional assessment of the shoulder. *Arthroscopy* 2003;19:1109–1120.
20. McClure P, Michener LA. Measures of adult shoulder function. *Arthritis Rheum* 2003;49:S50–S88.
21. 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.
22. Novak CB, Anastakis DJ, Beaton DE, Katz J. Patient-reported outcome after peripheral nerve injury. *J Hand Surg Am* 2009;34:281–287.
23. 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.
24. 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.
25. Roy JS, MacDermid JC, Woodhouse LJ. Measuring shoulder function: a systematic review of four questionnaires. *Arthritis Rheum* 2009;61:623–632.
26. Nunnally JC, Bernstein IH. Psychometric theory. 3rd ed. New York: McGraw Hill Inc; 1994.
27. McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? *Qual Life Res* 1995;4:293–307.
28. 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.
29. Lauchlan DT, McCaul JA, McCarron T. Neck dissection and the clinical appearance of post-operative shoulder disability: the post-operative role of physiotherapy. *Eur J Cancer Care (Engl)* 2008;17:542–548.
30. McNeely ML, Parliament MB, Seikaly H, et al. Effect of exercise on upper extremity pain and dysfunction in head and neck cancer survivors: a randomized controlled trial. *Cancer* 2008;113:214–222.
31. Kellerman SE, Herold J. Physician response to surveys. A review of the literature. *Am J Prev Med* 2001;20:61–67.
32. Leslie LL. Are high response rates essential to valid surveys? *Soc Sci Res* 1972;1:323–334.