Dysphagia 17:97–114 (2002)
DOI: 10.1007/s00455-001-0109-1

Dysphagia

The SWAL–QOL and SWAL–CARE Outcomes Tool for Oropharyngeal Dysphagia in Adults:
III. Documentation of Reliability and Validity

Colleen A. McHorney, PhD,[1] JoAnne Robbins, PhD,[2] Kevin Lomax, MHA,[3] John C. Rosenbek, PhD,[4] Kimberly Chignell, MS,[5] Amy E. Kramer, MS,[6] and D. Earl Bricker, MS[7]

[1]Roudebush VA Medical Center; Department of Internal Medicine and Geriatrics, Indiana University School of Medicine; Regenstrief Institute for Health Care; Indiana University Center for Aging Research, Indianapolis, Indiana; [2]William S. Middleton Memorial Veterans Hospital and University of Wisconsin–Madison Medical School, Madison, Wisconsin; [3]Ph.D. Program in Gerontology, Sanders–Brown Center on Aging, University of Kentucky, Louisville, Kentucky; [4]Department of Communicative Disorders, University of Florida, Gainesville, Florida; [5]William S. Middleton Memorial Veterans Hospital, Madison, Wisconsin; [6]Department of Surgery, University of Wisconsin–Madison Medical School, Madison, Wisconsin; [7]ACHOICE, Milwaukee, Wisconsin, USA

Submitted January 24, 2001; accepted August 6, 2001, with revision

Abstract. Advances in the measurement of swallowing physiologic parameters have been clinician-driven, as has the development of intervention techniques to modify swallowing pathophysiology. However, a critical element to determining the success of such efforts will be established by the patients themselves. We conceptualized, developed, and validated the SWAL–QOL, a 93-item quality-of-life and quality-of-care outcomes tool for dysphagia researchers and clinicians. With 93 items, the SWAL–QOL was too long for practical and routine use in clinical research and practice. We used an array of psychometric techniques to reduce the 93-item instrument into two patient-centered outcomes tools: (1) the SWAL–QOL, a 44-item tool that assesses ten quality-of-life concepts, and (2) the SWAL–CARE, a 15-item tool that assesses quality of care and patient satisfaction. All scales exhibit excellent internal-consistency reliability and short-term reproducibility. The scales differentiate normal swallowers from patients with oropharyngeal dysphagia and are sensitive to differences in the severity of dysphagia as clinically defined. It is intended that the standardization and publication of the SWAL–QOL and the SWAL–CARE will facilitate their use in clinical research and clinical practice to better understand treatment effectiveness as a critical step toward improving patients' quality of life and quality of care.

Key words: Deglutition — Deglutition disorders.

Over 30 years ago, Sidney Katz hypothesized that older individuals lose functions in the reverse order in which newborns gain them [1,2]. Thus, in models of human function [3–6], and in studies of functional status [1,3,4,7–15] and functional trajectories [4,12], the ability to feed is usually the last function to be lost (that is, it is the least prevalent of all disabilities affecting activities of daily living). Accordingly, dysphagia—the difficulty or inability to swallow—affects the most basic of sociobiological functions, the ability to eat and drink.

To date, standards of practice in the field of dysphagia largely reflect clinician-based assumptions about the impact of swallowing limitations on patients' health, function, and quality of life. While the advances in measurement of swallowing physiologic parameters have been clinician-driven [16–18], as has the development of intervention techniques to modify specific aspects of swallowing pathophysiology, a critical element in determining the success of such

---

Work performed at William S. Middleton Memorial Veterans Hospital, University of Wisconsin Hospitals and Clinics, Northwestern Memorial Hospital, Chicago, IL

Correspondence to: Colleen A. McHorney, Ph.D., Regenstrief Institute for Health Care, 1050 Wishard Blvd., RG, 6th Floor, Indianapolis, IN 46202, USA. Telephone: (317) 630-7664; fax: (317) 630-6611; E-mail: cmchorney@regenstrief.org

efforts will be established by the patients themselves. However, standardized instruments that measure outcomes of importance to patients and their families have been virtually nonexistent [19–21].

Toward this end (i.e., taking into explicit account the critical point of view of the patient [19–21]), we conceptualized, developed, and validated the SWAL–QOL, a 93-item quality-of-life and quality-of-care outcomes tool for dysphagia researchers and clinicians [22,23]. With 93 items, the SWAL–QOL is too long for practical and routine use in clinical research and practice and might prove too burdensome for dysphagic patients, who are typically older and chronically ill [22–24]. Furthermore, the SWAL–QOL is intended for use in clinical trials of treatment effectiveness [22]. To accommodate other data collection elements in efficacy and effectiveness studies, such as pathophysiologic and resource utilization outcomes, it is important to construct quality-of-life tools that are as brief as possible without sacrificing measurement precision.

With these considerations in mind, we used an array of psychometric techniques to reduce the 93-item instrument into two patient-centered outcomes tools: (1) the SWAL–QOL, a 44-item tool that assesses ten quality-of-life concepts, and (2) the SWAL–CARE, a 15-item tool that assesses quality of care and patient satisfaction. This article reports on item reduction techniques and extensive data on the reliability and validity of the SWAL–QOL and the SWAL–CARE.

Methods

Overview

Three phases of investigation were employed to develop and validate the SWAL–QOL and the SWAL–CARE. In Phase 1, we used focus groups to obtain qualitative data on quality-of-life and quality-of-care health outcomes from dysphagic patients and their caregivers. Using the focus group data, we wrote a large number of questionnaire items and constructed a pretest version of the SWAL–QOL [22]. In Phase 2, we pretested the SWAL–QOL and conducted preliminary psychometric tests [23]. The SWAL–QOL was revised to improve item quality and to reduce the length of the item pool. In Phase 3, we field tested the revised SWAL–QOL, further shortened it, and conducted thorough psychometric tests. We report the results of Phase 3 in this article.

Sampling

The sampling design and patient eligibility criteria for the SWAL–QOL study has been detailed extensively elsewhere [22] and so are described here only briefly. The three study phases used the same protocol for subject recruitment and sampling. Patients with mechanical or neurologic oropharyngeal dysphagia were recruited from outpatient speech, swallowing, neurology, and other medical clinics; ENT and other surgery clinics of VA hospitals; their affiliated university medical centers; and free-standing medical centers. In Phase 3, we recruited patients from Madison, WI, Chicago, IL, Little Rock, AR, Houston, TX, Atlanta, GA, and Long Beach, CA.

Patients were required to have had at least one videofluoroscopic swallowing examination demonstrating oropharyngeal dysphagia, as judged by an experienced speech-language pathologist, within three months of recruitment. We also required that they had a stable dysphagia, according to the patient and his or her clinician. If these two criteria were met, patients were invited to participate in the study. Patients were excluded if they were unable to provide informed consent, unable to speak and read English, refused to participate, or had evidence of active esophageal dysphagia.

We also recruited a total of 40 healthy male and female subjects between the ages of 45 and 95 years (average age = 72.9) as "normal" control subjects to test the ability of the SWAL–QOL to differentiate between those with and without a dysphagia. The control group was recruited from the general community, a hospital volunteer department, and an elder-hostel education program. None of these people reported prior or existing medical conditions and/or used medications that could potentially influence oropharyngeal motor performance or sensation. An oropharyngeal motor exam revealed normal structures and function of the bulbar musculature for all subjects. Most of this group (70%) was part of a larger study on the effects of aging on the biomechanics of swallowing [25].

Data Collection

The field survey consisted of a mail-out/mail-back questionnaire. It contained 123 items (93 SWAL–QOL items and another 30 items as validity criteria and demographic variables). We determined *a priori* that a sample of 400 patients would be sufficient to conduct psychometric tests. To achieve this, we implemented a mail survey protocol (which included altruistic, individualized, and personally signed cover letters [26]) that would yield a 75% response rate. Questionnaires that were not received within 14 days were followed up by a second mailing.

Data Analysis

Item Reduction

Reducing the length of the SWAL–QOL involved making a tradeoff between breadth and depth of coverage of quality-of-life and quality-of-care outcomes specific to dysphagia. At one extreme would be to use single-item indicators of multiple health concepts. At the other extreme would be to use long multi-item scales to measure only a few health concepts. Previous research has demonstrated that multi-item scales have superior content validity, are more reliable, achieve better breadth and depth of measurement, and yield less ceiling effects compared with single-item measures [27–31]. Thus, because of their clear psychometric merits, we constructed multi-item scales for each health concept. Because we did not want to sacrifice either breadth of measurement (comprehensiveness) or depth of measurement (precision), we decided to assess numerous health concepts using as few items as possible to construct each multi-item scale.

Each scale was constructed using Likert's method of summated ratings [32], which equally weighs each item and sums them into an overall scale score. All scales were linearly transformed to a 0-to-100 metric, with 100 indicating the most favorable state, 0 the least favorable, and scores in between representing the percentage of the total possible score achieved.

We used a variety of well-accepted item analysis techniques to reduce the length of the 93-item SWAL–QOL. We tested whether the quality-of-life and quality-of-care items formed a unidimensional structure by extracting principal components from the correlations among the 54 quality-of-life items and among the 39 quality-of-care items. Items that failed to correlate 0.40 [33] with the first principal component should be removed.

To achieve satisfactory reliability with short multi-item scales, it is necessary that each item correlate substantially with the scale it is hypothesized to represent. We used a correlation of 0.40 [34] (corrected for overlap) as the standard for supporting item-convergent validity. A correlation corrected for overlap is the correlation of an item with the sum of the other items in the same scale, thus removing the bias of correlating an item with itself. Tests of item-discriminant validity gauge the extent to which items correlate more highly with the concept they are hypothesized to represent than with different concepts. Hypothesized clusters of items are supported when correlations between items and their hypothesized scale are significantly higher (by two standard errors) than with other scales [35].

We used a two-parameter graded-response model [36,37] to evaluate item discrimination. The item discrimination parameter assesses the ability of an item to differentiate between people with different levels (or amounts) of the construct of interest. For example, if a mental health item received the same response from people with no depressive symptoms as it did from people with noteworthy depressive symptoms, the item would have low discrimination (i.e., it would fail to differentiate among levels of depression). Items with low discrimination parameters are reasonable candidates for removal because they yield little information about the trait being measured.

In practical terms, a scale score cannot be estimated with confidence if item-level missing data are high. Items with missing data rates greater than 10% were candidates for exclusion. Item frequency distributions were calculated and the range of responses for each item was assessed. Item distributions were examined for floor and ceiling effects, i.e., the respective clustering of subjects at the worst and best possible response categories. Floor and ceiling effects are undesirable because patients cannot improve on their scores if they are at the ceiling and cannot decline if they are at the floor. Items that had more than 15% of their responses at the floor, ceiling, or both [38] were candidates for exclusion.

Short-Form Instrument Validation

Scale Reliability. We computed Cronbach's α coefficient [39] to estimate the internal-consistency reliability of each scale. Reliability estimates greater than 0.80 are satisfactory for group-level research, while coefficients of 0.95 or greater are necessary for individual-patient decision-making [27,40]. We assessed test–retest reliability using Pearson's correlation coefficient and the intraclass correlation. Intraclass correlations of 0.75 or greater are considered excellent [41].

Features of Score Distributions and Scale–Scale Correlations. To determine the range of health states defined by each scale, we calculated the number of distinct levels of health enumerated by each scale, the observed score range, standard measures of central tendency, and the percent of the sample achieving the lowest (floor effect) and highest (ceiling effect) possible scores. Scale scores were correlated with each other to assess the extent to which they measured distinct aspects of swallowing sequelae.

Factorial Validity. We conducted a principal components analysis at the scale level to test the hypothesized dimensionality of the scales. Correlations between the scales and the first unrotated component test for the large dysphagia outcome factor hypothesized to be common to all scales. The pattern of correlations between the 13 scales and the rotated components test the validity of each scale relative to the hypothesized dimensionality. The hypothesized dimensions of the 13 scales are quality of life and quality of care. Within quality of life, two scales (sleep and fatigue) were written without any specific attribution to swallowing problems. Thus, we hypothesized that three factors would emerge: dysphagia-specific quality of life, generic quality of life, and quality of care.

Convergent and Discriminant Validity. Convergent validity is the extent to which a measure corresponds with other measures that assess similar concepts (attitudes or behaviors). Discriminant validity is the extent to which a measure does not correspond with conceptually distinct constructs. To assess scale-level convergent and discriminant validity, we included in the field test external measures of social functioning [42], general mental health [42,43], health perceptions [42,43], and loneliness [42] from the Medical Outcomes Study (MOS) and single-item measures of satisfaction with eating and drinking and overall satisfaction with care.

We hypothesized that (1) paired measures of the same underlying concept (e.g., generic and SWAL–QOL mental health) would correlate more highly with each other than with unpaired measures of different concepts (e.g., SWAL–QOL social function with generic health perceptions); (2) the psychosocial scales of the SWAL–QOL would be more highly correlated with each other than with SWAL–QOL scales measuring physical-health constructs (e.g., sleep and fatigue); and (3) measures of satisfaction with health would be more highly correlated with their paired domain than with unrelated domains. Consistent with past research [44–47], we hypothesized that the measures of quality of care and patient satisfaction would be positively, yet modestly, correlated with quality-of-life outcomes. We tested these hypotheses using product-moment correlation coefficients.

Clinical Validity. We used the method of known-groups validity [48,49] to test the clinical validity of the scales. To be useful in clinical applications, measures of patient-based health outcomes should meet two clinical validity standards: (1) discriminate between groups with and without disease and (2) distinguish between severity levels of a given condition [50]. These tests require the construction of mutually exclusive groups that differ in swallowing status based on external clinical criteria.

To test the validity of the SWAL–QOL to discriminate between patients with and without dysphagia, 40 healthy male and female subjects completed the symptom battery and the scales on food selection, eating desire, fear, communication, fatigue, and sleep. They did not complete the scales on mental health, social functioning, burden, and eating duration because these scales all had the words "swallowing problem" in the item stem, thus making it inapplicable to individuals without such problems. They also did not complete the quality-of-care and patient satisfaction scales

Table 1. Phase 3 validation recruitment

- 2654 patients screened
- Patients clinically ineligible (n = 1799)
  - —42% inpatient
  - —17% normal swallowing examination
  - —11% mental status
  - —7% inability to provide informed consent
  - —6% esophageal swallowing disorder
  - —4% patient denies dysphagia
  - —4% acute outpatients
  - —2% non-English speaking
  - —1% deceased
  - —6% other reason
- Patients clinically eligible (n = 855)
- Patients disqualified from participation (n = 215 or 25%)
  90% of patients could not be contacted
  10% of patients contacted but denied dysphagia
- Patients refused participation (n = 85 or 10%)
- Patients accepted participation (n = 555 or 65%)
- Patients participated (n = 386 or 70%)

Table 2. Descriptive characteristics of phase 3 validation sample (n = 386)

Characteristic	N	%
Age		
<45	23	6.0
45–54	52	13.5
55–64	87	22.5
65–74	110	28.5
75–84	95	24.6
85+	19	4.9
Mean age = 66.1 (13.17)		
Male	303	78.5
White	338	88.2
Primary reason for dysphagia		
Cancer	109	28.2
Vascular diseases	61	15.8
Degenerative neurologic diseases	49	12.7
Other neurologic diseases	36	9.3
Obstructive respiratory diseases	23	6.0
Trauma	15	3.9
Chronic medical conditions	17	4.4
Dementia	4	1.0
Other reasons	26	6.7
Unknown reason	46	11.9

because these healthy subjects would not have seen a swallowing clinician.

We hypothesized that the nondysphagic normal control group would score significantly higher (better health status) on each of the SWAL–QOL scales than patients with dysphagia. We tested this hypothesis using t-tests and general linear models, the latter to control for differences in age, sex, race, and education between dysphagia patients and the normal control group.

A second test of validity assessed whether scales are sensitive to differences in disease severity. The severity of dysphagia can be defined by its management strategy—whether an individual is on a restricted diet, must drink thickened liquids or none at all, or must use a feeding tube for nutrition. Patients answered items on the texture of their usual meals (regular diet, soft food diet, or pureed diet), the consistency of the liquids they drank (thin, thickened, or no liquid by mouth), and whether they took any of their nutrition by feeding tube (no or yes).

We also obtained information on the severity of a patient's dysphagia from the symptom battery. From these items, we constructed three scales measuring pharyngeal symptoms, oral symptoms, and saliva symptoms. We divided each symptom scale into quartiles and computed mean scores for the SWAL–QOL scales for each quartile. We hypothesized a linear and monotonic relationship between symptom severity and quality of life and quality of care. These hypotheses were tested using general linear models.

Results

Patient Characteristic

As shown in Table 1, 2654 patients across six sites were screened for participation in the field test, and two-thirds (67%) were clinically ineligible. The major reasons for clinical ineligibility were inpatient status (42%), normal swallowing examination (17%), and impaired mental status (11%). Of those who were eligible and agreed to participate, 70% returned their surveys. Table 2 shows demographic and clinical characteristics of the sample. The mean age was 66.1, and 58% were age 65 or older. Consistent with sampling at VA medical centers, over three-fourths were men. Caucasians constituted 90% of the sample. Reasons for the dysphagia were typical for those with oropharyngeal dysphagia (i.e., cancer, vascular disorders, and neurological disorders).

Item Reduction

The 54 items measuring quality of life were subjected to a principal components analysis (data available upon request). The eigenvalue for the first unrotated factor was 25.3 and accounted for 47% of the variance. All items loaded highly on the first principal component (>0.44). For the 39 items measuring quality of information and patient satisfaction, the eigenvalue for the first unrotated factor was 13.5 and accounted for 61% of the variance. All items loaded highly on the first principal component (>0.73).

Table 3. Summary of rationale for item deletion

	High missing data	Unacceptable skewness	Poor discriminant validity	Poor item discrimination	Awkward phrasing
Burden		Burd2	Burd2	Burd2	
Food selection			Food4	Food4	
Eating loss			Eat2, 13, 14		
Eating duration		Eat4	Eat4	Eat4	Eat19
Eating desire		Eat9			
Fear		Fear6	Fear6	Fear6	
Sleep					Sleep7
Fatigue	Fat6	Fat1			Fat1
Communication	Com6	Com3, 6, 7		Com3	
Mental health		MH3, 9, 12, 17, 23		MH3, 17, 23	
Social	Soc12	Soc4, 7	Soc12	Soc4, 12	Soc4, 8
Advice	Adv8, 9	Adv9, 13	Adv9	Adv9, 13	
Patient satisfaction	Sat21	Sat3, 21, 22		Sat3, 21	
Symptoms				Sym6, 10, 11	

The rationale for item reduction is summarized in Table 3. Most of the items were removed because of problems with skewness (ceiling effects), item discrimination, and discriminant validity. The eating loss scale was dropped altogether because all of its items failed tests of item discriminant validity. Two domains were hypothesized for mental health (self-image and psychological distress), social functioning (social and role functioning), advice (intake and techniques advice), and patient satisfaction (technical and interpersonal). The data, however, did not support any of these dual domains, so a single scale was constructed for mental health, social functioning, clinical advice, and patient satisfaction.

We reduced the SWAL–QOL by 36%, from 93 items to 59 items (verbatim item content is presented in Table 4). At this point in tool development, we decided to separate the items on quality of information and patient satisfaction into a different instrument called the SWAL–CARE. Thus, the SWAL–QOL has 44 items and the SWAL–CARE has 15. We made this decision because it might take more than one visit for patients to be able to reliably rate their swallowing clinicians on quality of information and advice and satisfaction with their care, whereas the SWAL–QOL could be completed on or before a first visit.

Respondent Burden

Administration time for the 123-item field survey (93 SWAL–QOL items, and 30 items as validity criteria and demographics) averaged about 3.2 items per minute. It should take respondents an average of 14 minutes to complete the 44-item SWAL–QOL and about 5 minutes to complete the 15-item SWAL–CARE.

Reliability

Reliability estimates are presented in Table 5. Only one scale (fear, $\alpha = 0.79$) failed to attain the reliability standard of 0.80 recommended for group-level research. Otherwise, five scales had α coefficients from 0.80 to 0.89, and seven scales had $\alpha > 0.90$. Two-week retest reliability estimates were all in the substantial to excellent range [41,51]. Pearson correlations ranged from 0.60 to 0.91, with a median of 0.76. Intraclass correlations ranged from 0.59 to 0.91, with a median of 0.76. Thus, these scales have excellent internal-consistency reliability and short-term stability.

Features of Score Distributions

As shown in Table 6, the full range of the 0-to-100 score distribution was observed for all scales. Floor effects ranged from 2.4% (fear) to 19.1% (eating duration). Ceiling effects ranged from 4.2% (fatigue) to 42.2% (patient satisfaction).

Scale–Scale Correlations

Table 7 presents scale–scale correlations. Correlations ranged from 0.05 to 0.78. Scales measuring similar constructs correlated more highly with each other than those measuring different constructs. For example, the eating scales (burden, food selection,

Table 4. Abbreviated item content[a]

Burden
Dealing with my SP is very difficult
SP is a major distraction in my life

Eating Duration
It takes me longer to eat than other people
It takes me forever to eat a meal

Eating Desire
Most days, I don't care if I eat or not
I don't enjoy eating anymore
I'm rarely hungry anymore

Symptom Frequency
Coughing
Choking when you eat food
Choking when you take liquids
Having thick saliva or phlegm
Gagging
Having excess saliva or phlegm
Having to clear your throat
Drooling
Problems chewing
Food sticking in your throat
Food sticking in your mouth
Food/liquid dribbling out your mouth
Food/liquid coming out your nose
Coughing food/liquid out your mouth

Food Selection
Figuring out what I can eat is a problem for me
It is difficult to find foods I both like and can eat

Communication
People have a hard time understanding me
It's been difficult for me to speak clearly

Fear
I fear I may start choking when I eat food
I worry about getting pneumonia
I am afraid of choking when I drink liquids
I never know when I am going to choke

Mental Health
My SP depresses me
I get impatient dealing with my SP
Being so careful when I eat or drink annoys me
My SP frustrates me
I've been discouraged by my SP

Social
I do not go out to eat because of my SP
My SP makes it hard to have a social life
My usual activities have changed BOM SP
Social gatherings are not enjoyable BOM SP
My role with family/friends has changed BOM SP

Fatigue
Feel exhausted
Feel weak
Feel tired

Sleep
Have trouble falling asleep
Have trouble staying asleep

Advice (SWAL–CARE)
Food I should eat
Foods I should avoid
Liquids I should drink
Liquids I should avoid
Techniques to help get food down
Techniques to help me avoid choking
When I should contact a swallowing clinician
My treatment options
What to do if I start to choke
Signs that I am not getting enough to eat or drink
Goals of the treatment for my SP

Patient Satisfaction (SWAL–CARE)
Had confidence in your swallowing clinicians
Swallowing clinicians explained treatment to you
Swallowing clinicians spent enough time with you
Swallowing clinicians put your needs first

[a] BOM = because of my; SP = swallowing problem. Item content is abbreviated. The SWAL–QOL and SWAL–CARE are available free of charge upon request.

eating duration, and eating desire) correlated, on average, 0.51 with each other. This is evidence of convergent validity. The weakest correlations were between the quality-of-life and quality-of-care scales (information, advice, and satisfaction). These correlations ranged from 0.05 to 0.33. This is evidence of discriminant validity.

Factor Structure

As shown in Table 8, whether derived by varimax or oblique rotation, three distinct factors emerged. The first factor (eigenvalue = 5.58, accounting for 43% of the variance) is composed of the dysphagia-specific quality-of-life scales. The second factor (eigenvaluae = 2.03, accounting for an additional 16% of the variance) is composed of the quality-of-care and patient-satisfaction scales. The third factor (eigenvalue = 1.02, accounting for an additional 8% of the variance) is composed of sleep and fatigue, which were not written specific to swallowing disorders.

These factor loadings define the contribution of each scale to the underlying construct of dysphagia health outcomes. Thus, food selection, burden, mental health, and social functioning are the primary

Table 5. Reliability estimates

Scale	Reliability estimates		
		Test–retest^b	
	Internal consistency^a	Pearson	Intraclass
SWAL–QOL			
Food selection	0.89	0.83	0.83
Burden	0.89	0.60	0.59
Mental health	0.94	0.80	0.80
Social functioning	0.94	0.88	0.89
Fear	0.79	0.74	0.74
Eating duration	0.80	0.64	0.64
Eating desire	0.86	0.91	0.91
Communication	0.91	0.76	0.76
Sleep	0.81	0.80	0.81
Fatigue	0.90	0.85	0.85
SWAL–CARE			
Clinical information	0.95	0.69	0.69
General advice	0.94	0.63	0.64
Patient satisfaction	0.92	0.62	0.62

^{a}Internal-consistency sample size = 386.
^{b}Test–retest sample size = 37; average test–retest interval = 15.5 days.

Table 6. Features of score distributions (n = 386)

	No. Items	No. Levels	Range	Mean	Median	SD	% Floor	% Ceiling
SWAL–QOL								
Food selection	2	9	0–100	55.5	63.0	33.6	10.8	18.8
Burden	2	9	0–100	46.0	50.0	31.3	16.0	8.0
Mental health	5	21	0–100	52.4	52.5	30.7	6.1	8.5
Social functioning	5	21	0–100	60.5	65.0	32.8	5.0	21.5
Fear	4	17	0–100	59.3	63.0	27.1	2.4	9.1
Eating duration	2	9	0–100	45.0	38.0	34.0	19.1	11.6
Eating desire	3	13	0–100	68.0	75.0	32.0	4.1	32.4
Communication	2	9	0–100	56.5	50.0	31.4	6.8	18.9
Sleep	2	9	0–100	58.4	63.0	30.2	5.8	17.5
Fatigue	3	13	0–100	46.7	50.0	28.6	12.4	4.2
SWAL–CARE								
Clinical information	6	31	0–100	54.9	60.0	26.1	3.8	7.1
General advice	5	26	0–100	46.3	44.0	28.4	7.8	6.5
Patient satisfaction	4	13	0–100	78.5	92.0	27.0	3.1	42.2

contributors to dysphagia-specific quality of life, followed by fear, eating duration, eating desire, and communication. Sleep and fatigue contribute equally to generic quality of life. Advice contributes slightly more to quality of care than does patient satisfaction. All tables that follow present the SWAL–QOL and the SWAL–CARE in this order to facilitate interpretation.

Convergent and Discriminant Validity

Table 9 presents correlation coefficients between the SWAL–QOL and the SWAL–CARE scales and the health perceptions [42,43], social functioning [42], loneliness [34], and general mental health [42,43] measures from the MOS. The correlations between SWAL–QOL social functioning scale and MOS social functioning and loneliness measures are large (r = 0.50 and r = 0.56, respectively) as is the correlation between SWAL–QOL mental health and MOS mental health items (r = 0.51). This is evidence of convergent validity—there is good agreement between different measures (generic vs. disease-specific) of the same construct. Furthermore, each matched correlation (SWAL–QOL social with MOS social and loneliness; SWAL–QOL mental health with MOS

Table 7. Scale–scale correlations ($n = 386$)[a]

	SWAL–QOL										SWAL–CARE		
	Food	Burd	MH	Social	Fear	Eatdur	Eatdes	Com	Sleep	Fatig	Clin	Gen	PatSat
Food		0.66	0.68	0.66	0.53	0.44	0.53	0.37	0.32	0.36	0.22	0.22	0.26
Burd			0.69	0.62	0.55	0.43	0.46	0.38	0.31	0.37	0.26	0.31	0.29
MH				0.74	0.64	0.47	0.58	0.49	0.42	0.47	0.25	0.33	0.23
Social					0.52	0.47	0.62	0.51	0.41	0.39	0.17	0.22	0.22
Fear						0.35	0.37	0.40	0.39	0.41	0.23	0.27	0.24
Eatdur							0.52	0.33	0.19	0.33	0.18	0.28	0.21
Eatdes								0.36	0.36	0.45	0.21	0.27	0.24
Com									0.29	0.31	0.15	0.15	0.16
Sleep										0.53	0.05[b]	0.07[b]	0.07[b]
Fatig											0.23	0.25	0.24
Clin												0.78	0.64
Gen													0.63

[a] Food = food selection; Burd = burden; MH = mental health; Social = social functioning; Fear = fear; Eatdur = eating duration; Eatdes = eating desire; Com = communication; Sleep = sleep; Fatig = fatigue; Clin = clinical information; Ger = general advice; PatSat = patient satisfaction.

[b] Correlations are *not* statistically significant at $p < 0.05$.

Table 8. Scale-level factor loadings

	Varimax rotation			Oblique rotation		
	Factor 1 Dysphagic-specific QOL	Factor 2 Quality of care	Factor 3 Generic QOL	Factor 1 Dysphagic-specific QOL	Factor 2 Quality of care	Factor 3 Generic QOL
Food selection	0.83			0.91		
Burden	0.81			0.88		
Mental health	0.81			0.81		
Social functioning	0.82			0.80		
Fear	0.71			0.71		
Eating duration	0.59			0.58		
Eating desire	0.58			0.50		
Communication	0.54			0.54		
Sleep			0.83			0.87
Fatigue			0.81			0.83
Clinical information		0.89			0.91	
General advice		0.88			0.89	
Patient satisfaction		0.84			0.85	

mental health) is higher than all the other correlations in each row (except for fatigue and MOS mental health). This provides further evidence of convergent validity. We also hypothesized that measures of quality of life would correlate more highly with one another than with quality of care. This, too, is borne out by the data: The average correlation between the generic measures and the SWAL–QOL scales is 0.35, while the average correlation between the generic measures and the SWAL–CARE scales is 0.19. This provides evidence of discriminant validity.

We correlated the SWAL–QOL and the SWAL–CARE scales with single-item measures of subjects' satisfaction with eating and drinking and an overall rating of the quality of their care. As shown in Table 10, the correlation between satisfaction with eating and burden is 0.57 and that with food selection is 0.56. Eating satisfaction also correlates substantially with social functioning ($r = 0.55$), mental health ($r = 0.52$), and eating desire ($r = 0.49$). This is further evidence of convergent validity. Satisfaction with drinking is most highly correlated with fear ($r = 0.41$), social functioning ($r = 0.40$), and burden ($r = 0.40$). As hypothesized, the overall quality-of-care rating is highly correlated with the three measures of quality of care (average correlations = 0.74)

Table 9. Correlations with MOS measures ($n = 386$)[a]

	SWAL–QOL										SWAL–CARE		
MOS measures	Food	Burd	MH	Social	Fear	Eatdur	Eatdes	Com	Sleep	Fatig	Clin	Gen	PatSat
Health perceptions	0.15	0.11	0.20	0.20	0.26	0.11	0.17	0.20	0.29	0.50	0.16	0.19	0.18
Social function	0.35	0.31	0.40	0.50	0.37	0.24	0.32	0.36	0.39	0.49	0.10	0.17	0.19
Loneliness	0.42	0.36	0.52	0.56	0.43	0.29	0.42	0.36	0.44	0.48	0.18	0.23	0.27
Mental health	0.38	0.34	0.51	0.46	0.43	0.20	0.40	0.28	0.43	0.52	0.15	0.20	0.22

[a] All correlations significant at $p < 0.05$.

Table 10. Correlations with satisfaction ratings ($n = 386$)

	SWAL–QOL										SWAL–CARE		
	Food	Burd	MH	Social	Fear	Eatdur	Eatdes	Com	Sleep	Fatig	Clin	Gen	PatSat
Satisfaction with eating	0.56	0.57	0.52	0.55	0.35	0.34	0.49	0.29	0.28	0.35	0.32	0.37	0.33
Satisfaction with drinking	0.37	0.40	0.36	0.40	0.41	0.15	0.26	0.26	0.22	0.29	0.33	0.28	0.28
Overall quality of care	0.22	0.27	0.23	0.19	0.16	0.12	0.15	0.13	0.08[a]	0.24	0.71	0.71	0.79

[a] Correlations are *not* significant at $p < 0.05$.

and is weakly correlated with the SWAL–QOL scales (average correlation = 0.18). This is further evidence of discriminant validity.

Clinical Validity: Dysphagia Patients vs. Normal Swallowers

Table 11 presents the results of clinical validity for each individual SWAL–QOL symptom. As hypothesized, dysphagic patients scored lower (worse) on all of the symptoms. The greatest differences were observed for food sticking in throat, having to clear one's throat, coughing when food is stuck, choking on food, having excess saliva, and gagging. On average, the dysphagic patients reported that these symptoms happened ''sometimes'' while the normal swallowers reported that they occurred ''almost never.''

Table 12 presents the results of the differences between normal swallowers and dysphagic patients on the SWAL–QOL scales and the MOS scales. As hypothesized, statistically significant differences were observed across all the scales. The greatest difference [well over one standard deviation ($z = 1.31$)] was observed for the fear scale: the average score for dysphagics was 59 while that for normal swallowers was 96, approaching the ceiling. Almost a full standard deviation difference was observed between the two groups on communication (56 vs. 88). Large and statistically significant differences were also observed for eating desire (68 vs. 93), food selection (55 vs. 77), sleep (58 vs. 76), and fatigue (47 vs. 73). Differences between the two groups on MOS measures were also very large and statistically significant. On average, people with dysphagia rated their health to be between ''fair and good,'' while the normal swallowers reported theirs to be between ''good and very good.'' The normal swallowers reported almost perfect social functioning (92) while that for dysphagics was nearly one standard deviation lower (59). Based on the MHI-5's cutoff point of 52 [52,53], 33% of these dysphagic patients would screen positive for major depression, while only 5% of the normal swallowers would. All of these results remained large and statistically significant when scores adjusted for age, sex, race, and education were compared.

Clinical Validity: Dysphagic Patients with Oral Nutrition vs. Tube Feeding

Six of the 14 tests involving tube-feeding status were significant at the 0.05 level and three were significant at the 0.08 level (Table 13). The largest differences observed between tube feeders and regular eaters were for social functioning (34 vs. 65, z difference of 0.93) and burden (25 vs. 49, z difference of 0.80). Large and statistically significant differences were also observed for food selection, mental health, fear, and eating desire, with the tube feeders all scoring lower. The SWAL–CARE scales were not sensitive to tube-feeding status.

Table 11. Results of clinical validity: differences between "normal" swallowers and dysphagics

Symptom	Dysphagic (n = 386)	Normals (n = 40)	Mean difference	p value
Food stick throat	2.91	4.55	1.64	0.000
Clear throat	2.38	3.77	1.39	0.000
Cough-food stuck	3.39	4.72	1.33	0.000
Choke on food	3.18	4.47	1.29	0.000
Excess saliva	2.86	4.15	1.29	0.000
Gagging	3.48	4.74	1.26	0.000
Thick saliva	2.65	3.85	1.20	0.000
Choke on liquid	3.38	4.55	1.17	0.000
Food stick mouth	3.56	4.72	1.16	0.000
Problem chewing	3.50	4.64	1.14	0.000
Coughing	2.55	3.60	1.05	0.000
Drooling	3.57	4.59	1.02	0.000
Dribble from mouth	3.80	4.74	0.94	0.000
Dribble from nose	4.24	4.87	0.63	0.000

Table 12. Results of clinical validity: differences between "normal" Swallowers and dysphagics

Measures	Dysphagics (n = 386)	Normals (n = 40)	z difference	p value
SWAL–QOL scales				
Food selection	55.5	76.6	0.63	0.000
Fear	59.3	96.0	1.31	0.000
Eating desire	68.0	93.2	0.80	0.000
Communication	56.5	87.8	0.99	0.000
Sleep	58.4	76.3	0.59	0.000
Fatigue	46.7	73.0	0.92	0.000
MOS measures				
Health perceptions	38.4	67.5	1.09	0.000
Social functioning	59.1	92.0	0.95	0.000
Mental health	61.4	78.1	0.71	0.000

All of these findings were maintained when the scores were adjusted for age, education, sex, and race.

Clinical Validity: Dysphagic Patients Differing in Food Texture

The greatest differences observed among the three oral-intake groups were in social functioning, eating duration, food selection, mental health, and burden (Table 14). Only two comparisons were not statistically significant: clinical information and general advice. Otherwise, the data confirmed the hypothesis that dysphagia severity, as reflected by food texture, is associated with worse quality of life. The scores for patients on pureed diets are particularly low. The three food texture groups also differed significantly in symptom severity (data not shown). For 11 of the 14 symptoms (all but choke on liquid, excess saliva, and drooling), individuals with regular diets had the least symptom severity compared with those with soft or pureed diets.

Clinical Validity: Dysphagic Patients Differing in Liquid Consistency

Only two significant differences in the SWAL–QOL scales were observed among the groups that differed in the consistency of their liquids (Table 15). Patients drinking thickened liquids scored lower on burden and social functioning than patients with thin liquids, and patients taking no liquid by mouth scored lower than those taking thickened liquids.

Clinical Validity: Symptom Severity and Quality of Life

The prevalence of the 14 symptoms is shown in Table 16. Clearing one's throat was the most prevalent symptom, followed by coughing and having thick saliva. Drooling and having food/liquid dribble from the mouth and nose were the least prevalent symptoms. To achieve parsimony in data analysis and interpretation, we constructed three symptom severity

Table 13. Results of clinical validity: differences by tube-feeding status

	Tube feeders ($n = 57$)	Non-tube feeders ($n = 323$)	z difference	p value
SWAL–QOL				
Food selection	36.1	58.7	0.67	0 0000
Burden	24.5	49.5	0.80	0.0000
Mental health	38.8	55.0	0.53	0.0001
Social functioning	34.6	65.1	0.93	0.0000
Fear	49.1	60.9	0.44	0.003
Eating duration	36.6	45.9	0.27	0.074
Eating desire	55.2	70.0	0.46	0.002
Communication	49.3	57.4	0.26	0.076
Sleep	51.3	59.5	0.27	0.062
Fatigue	44.6	46.7	0.07	0.616
SWAL–CARE				
Clinical information	59.2	53.9	−0.20	0.185
General advice	47.7	45.9	−0.06	0.680
Patient satisfaction	74.9	79.1	0.15	0.305

Table 14. Results of clinical validity[a] by food texture of regular diet

	Pureed food ($n = 20$)	Soft food ($n = 103$)	Regular food ($n = 214$)	p value
SWAL–QOL				
Food selection	26.4	35.3	70.1	0.0001
Burden	21.4	34.4	58.2	0.0001
Mental health	25.2	39.8	63.6	0.0001
Social functioning	29.2	43.5	77.1	0.0001
Fear	53.0	52.0	65.0	0.0001
Eating duration	10.0	26.5	56.4	0.0001
Eating desire	50.8	50.3	79.4	0.0001
Communication	37.7	46.3	65.3	0.0001
Sleep	47.1	50.9	64.5	0.0001
Fatigue	32.5	38.0	52.1	0.0001
SWAL–CARE				
Clinical information	56.8	50.4	57.4	0.12
General advice	40.9	41.0	49.2	0.08
Patient satisfaction	65.0	73.9	82.5	0.005

[a] Sample limited to persons taking oral nutrition.

scales. The 14 symptoms were subjected to a principal components analysis and they all loaded highly on the first unrotated principal component (0.43 to 0.73, an average r of 0.60).

To better interpret the scales, we used orthogonal rotation. One symptom (food/liquid dribble from nose) did not load on any of the three rotated factors, while another (food sticking in mouth) had equal and large correlations with two factors. These two symptoms were not included in any of the symptom scales. Three factors were then derived: (1) a seven-item factor on pharyngeal symptoms (coughing, choking on food, choking on liquids, gagging, clearing throat, food sticking in throat, and coughing when food stuck); (2) a two-item factor on saliva symptoms (thick saliva and excess saliva); and (3) a three-item factor on oral symptoms (drooling, problem chewing, and food/liquid dribble from the mouth). Similar results were observed with oblique rotation. The Cronbach's α was 0.85 for pharyngeal symptoms, 0.78 for saliva symptoms, and 0.72 for oral symptoms.

To understand the association between symptom severity and quality of life and quality of care, we divided each symptom scale into quartiles. As shown in Table 17, all of the scales were sensitive to gradients in pharyngeal symptom status, and the greatest differences were observed for fear, food selection, burden, and mental health. The smallest differences, albeit statistically significant,

Table 15. Results of clinical validity by liquid consistency

	No liquid by mouth (n = 17)	Thickened liquid (n = 48)	Thin liquid (n = 313)	p value
SWAL–QOL				
Food selection	47.1	49.9	56.9	0.260
Burden	22.9	37.7	48.7	0.000
Mental health	47.5	50.5	53.2	0.675
Social functioning	37.8	49.6	63.4	0.000
Fear	49.5	64.1	59.1	0.183
Eating duration	54.1	35.0	45.6	0.079
Eating desire	75.6	60.0	68.5	0.152
Communication	55.2	49.7	57.4	0.282
Sleep	43.5	60.3	58.9	0.113
Fatigue	39.2	50.5	46.2	0.350
SWAL–CARE				
Clinical information	63.0	53.0	54.7	0.451
General advice	54.0	40.0	46.9	0.199
Patient satisfaction	75.4	73.7	79.4	0.416

Table 16. Prevalence of dysphagia symptoms (n = 386)

	Almost always (%)	Often (%)	Sometimes (%)	Hardly ever (%)	Never (%)
Clear throat	23	35	28	8	6
Coughing	18	30	35	13	4
Thick saliva	22	28	24	13	12
Excess saliva	19	26	22	14	18
Food stick throat	15	22	34	14	15
Choke on food	7	16	41	23	13
Choke on liquid	6	14	36	25	20
Cough-food stick	7	16	32	19	25
Gagging	6	12	31	28	22
Problem chewing	14	10	20	20	35
Food stick mouth	9	11	22	28	29
Drooling	8	15	24	17	35
Dribble from mouth	7	9	20	25	38
Dribble from nose	2	5	15	24	54

were observed for information and patient satisfaction. Similar results were observed for the oral symptom severity (Table 18), although the gradients were somewhat less dramatic. Differences in quality of life by saliva symptom severity (Table 19) were also statistically significant, although less pronounced than pharyngeal and oral symptoms. For oral and saliva symptom severity, the two information scales did not meet the 0.05 level of statistical significance.

Discussion

Considerable progress has been made in measuring the biomechanical signs of dysphagia [16–18] and in standardizing its evaluation and treatment [54]. The value of these advances, however, has been incompletely understood because outcome assessments have been infrequent and largely limited to physiologic outcomes. The challenge for dysphagia researchers and clinicians is to continue making measurement advances toward a synergistic and thorough compendium of health outcomes tools specifically for dysphagia that represent a balance between clinician-driven parameters and those that represent the patient's perspective. As a step toward achieving this end, we constructed the SWAL–QOL and the SWAL–CARE outcomes tools and reported herein extensive and strong evidence of their reliability and validity.

Validity assessment is a never-ending process, and validity is always inferred, never established [55]. As the SWAL–QOL and the SWAL–CARE become more widely used, additional evidence about

Table 17. Results of clinical validity by pharyngeal[a] symptom severity

	Q1[b] (most symptomatic)	Q2[b]	Q3[b]	Q4[b] (least symptomatic)	*p* value
SWAL–QOL					
Food selection	35.5	44.0	72.2	75.8	0.0001
Burden	23.3	37.6	59.8	67.0	0.0001
Mental health	31.0	42.6	70.0	72.3	0.0001
Social functioning	43.2	50.3	74.9	75.7	0.0001
Fear	36.5	54.0	68.4	81.2	0.0001
Eating duration	31.0	34.7	51.1	61.3	0.0001
Eating desire	57.8	61.0	78.3	79.6	0.0001
Communication	42.6	54.4	66.2	67.2	0.0001
Sleep	45.5	58.4	64.6	67.9	0.0001
Fatigue	32.7	43.4	53.5	59.7	0.0001
SWAL–CARE					
Clinical information	49.6	52.8	57.5	63.2	0.02
General advice	39.0	41.0	50.6	55.3	0.003
Patient satisfaction	73.6	75.6	86.3	83.3	0.01

[a] Aggregation of coughing, choking, gagging, clear throat, food sticking throat, and coughing-stick.
[b] Q = quartile.

Table 18. Results of clinical validity by oral[a] symptom severity

	Q1[b] (most symptomatic)	Q2[b]	Q3[b]	Q4[b] (least symptomatic)	*p* value
SWAL–QOL					
Food selection	35.1	58.0	57.1	72.4	0.0001
Burden	32.7	39.8	53.2	56.3	0.0001
Mental health	35.6	49.6	55.6	66.3	0.0001
Social functioning	36.9	60.7	65.3	77.7	0.0001
Fear	47.9	58.6	61.6	71.3	0.0001
Eating duration	27.5	33.4	49.3	60.9	0.0001
Eating desire	50.5	67.3	70.9	81.0	0.0001
Communication	34.0	51.7	63.7	76.2	0.0001
Sleep	44.0	57.0	61.8	69.2	0.0001
Fatigue	34.0	45.4	49.2	57.8	0.0001
SWAL–CARE					
Clinical information	53.3	49.4	56.3	60.8	0.09
General advice	42.6	43.5	46.7	54.2	0.09
Patient satisfaction	74.2	74.3	81.2	85.0	0.03

[a] Aggregation of drooling, problems chewing, and dribble out of mouth.
[b] Q = quartile.

their reliability, validity, and responsiveness will accumulate. The original SWAL–QOL project was unable to support tests of responsiveness (sensitivity to change over time) due to limitations in time and funding. Such future tests will be important to document the sensitivity of the scales to treatment effects and natural history. That the SWAL–QOL scales were sensitive to disease severity as defined by food texture, symptom severity, and tube feeding provide indirect support for their ability to be responsive in longitudinal studies. A recent case study reported SWAL–QOL scores pre- and post-behavioral intervention for a dysphagic patient who had undergone surgery and radiation for head and neck cancer [56]. After eight weeks of a rigorous lingual exercise program, there were substantial changes in the expected direction in seven of the 13 scales (food selection, fear, sleep, fatigue, communication, mental health, and social functioning). Even more dramatic improvement was observed after a second 8-week intervention on all of these scales, and improvements emerged for eating duration and eating desire after this second 8-week intervention period.

With one exception, estimates of internal-consistency reliability for each scale exceeded the

Table 19. Results of clinical validity by saliva[a] symptom severity

	Q1[b] (most symptomatic)	Q2[b]	Q3[b]	Q4[b] (least symptomatic)	*p* value
SWAL–QOL					
Food selection	44.4	57.3	55.5	65.4	0.0001
Burden	32.6	44.4	49.3	59.0	0.0001
Mental health	40.0	48.6	56.7	65.9	0.0001
Social functioning	46.3	56.0	64.2	76.4	0.0001
Fear	47.7	60.4	63.5	67.9	0.0001
Eating duration	36.7	40.0	50.5	54.3	0.0001
Eating desire	55.4	64.4	75.2	77.9	0.0001
Communication	45.2	55.5	56.2	70.2	0.0001
Sleep	46.7	59.2	59.5	67.1	0.0001
Fatigue	35.1	46.9	48.2	57.1	0.0001
SWAL–CARE					
Clinical information	53.4	50.7	56.5	60.7	0.09
General advice	41.7	44.5	47.0	52.6	0.16
Patient satisfaction	71.3	80.9	77.4	83.5	0.04

[a] Aggregation of thick saliva and excess saliva.
[b] Q = quartile.

recommended standards of 0.80 for group comparisons. Since the fear scale at 0.79 missed the standard by only 0.01, we feel confident in recommending it for group-level research. Some of the scales (mental health, social functioning, clinical information, and general advice) had internal-consistency reliability that met or closely approached the standard of 0.95 for individual-patient assessment.

Floor effects exceeded 15% for only two scales: burden (16%) and eating duration (19%). Such floor effects are emblematic of the everyday hardships of dysphagia in that it changes a person's life in many ways, and the length of time it takes to eat a meal is more than a mere nuisance. Ceiling effects were large for the patient satisfaction scale (42%) and eating desire (32%). A high ceiling effect for patient satisfaction is consistent with that reported for other patient satisfaction measures [57–60]. In less-healthy samples, these ceiling effects will attenuate, while they will increase in healthier samples. Since our sample of dysphagics was on the healthier side, as defined by food texture, liquid consistency, and use of feeding tubes, future users of the SWAL–QOL and the SWAL–CARE should observe lower ceiling effects. Overall, the floor and ceiling effects observed here should not prove problematic for studies of treatment effectiveness since ample room is available for patients to both improve and decline on the SWAL–QOL and the SWAL–CARE scales.

The lack of strong covariance between quality of life, as defined by the SWAL–QOL scales, and quality of care, as defined by the SWAL–CARE scales, should not cast doubt on the validity of any of these measures since past research has shown health status and patient satisfaction to be marginally related [44]. Patient satisfaction is an important component of a comprehensive outcomes assessment [61,62] because it is associated with improved patient compliance [63–65], less physician switching [57,66,67] and plan disenrollment [67–73], less out-of-plan utilization [74], and greater intention to remain with a provider [63,75,76]. We created separate tools to assess quality of life and quality of care because we recognized that one or the other tool may be more or less appropriate depending on the application and population. Thus, each investigator should determine for him- or herself whether only one or both tools are clinically or theoretically relevant to the specific question at hand.

Using the MOS general mental health measure and its cutoff point of 52, we found that 33% of our dysphagia patients met clinical criteria for major depression. Such a high rate of mental health disability is of great clinical significance since mental health problems are underdetected [77,78] and undertreated [79] and result in excess utilization [80,81] and reduced quality of life [82]. Such a high rate of mental disability underscores the importance of measuring disease-specific mental health in the SWAL–QOL. Future work on the SWAL–QOL will establish a sensitive and specific cutoff point on the SWAL–QOL mental health scale for screening for major depression. Armed with such an interpretational benchmark, dysphagia clinicians could quickly screen patients for mental disability and make appropriate specialty referrals if patients screen positive.

Including a normal control group in the SWAL–QOL project provided a preliminary cornerstone for score interpretation. As hypothesized, normal swallowers experience virtually no fear of eating or drinking (average score = 96) and negligible perturbations in their eating desire (average score = 93). They report minimal impediments in communication (average score = 88), but a smattering of problems with food selection, sleep, and fatigue. An important next step for the SWAL–QOL and the SWAL–CARE is to collect normative data on each scale to assist investigators in score interpretation.

The validity results by dysphagia severity support the clinical validity of the scales. People taking nutrition by tube feeding had lower scores on the SWAL–QOL scales than regular feeders. The greatest differences were observed for social functioning and burden, which are clinically and socially intuitive. At the same time, people on tube feeding often remarked in the margins of the survey that questions specific to tube feeding need to be developed. One patient wrote: "You should ask more questions about how people feel and live on a G tube." Another wrote: "You should word your questions so that people like myself who live with a G tube can answer a little better." Future work will seek to develop and validate modules for persons taking nutrition nonorally.

Dysphagia severity, as reflected by food texture, had a noteworthy impact on the SWAL–QOL scales. Patients with pureed diets had scores approaching the floor of many of the scales. In secondary analyses, we found that 33% of patients with pureed diets screened positive for major depression and 35% rated their health as poor. Poor self-ratings of health have been associated with an increased risk of mortality [83]. Clinicians need to be aware of the subjective hardships that seem to result from a pureed diet. In fact, these findings may be interpreted as motivation for clinicians to provide alternative treatments to diet modification (such as postural adjustments and maneuvers) for the more severely impaired dysphagic patients, implementing pureed diet as a "last resort." Most important, perhaps, with the SWAL–QOL and the SWAL–CARE, dysphagia-specific quality-of-life outcomes can be used in studies comparing behavioral interventions to diet modification, thus providing clinicians and payers with reliable and valid patient-centered outcomes to influence clinical decision-making and reimbursement policies.

Only the burden and social functioning scale differentiated patients by consistency of liquid intake. Most of the patients in the sample (83%) were drinking thin liquids; thus, the restricted variance in liquid consistency may have attenuated expected differences between the groups. The lack of significant findings for some scales may also be explained by item content. For instance, the items assessing food selection, eating desire, and eating duration are focused on food, not liquid, and might be expected to be insensitive to limitations in fluid intake. With few exceptions, most of the observed results for liquid consistency were in the expected direction.

The 14-item symptom battery should prove useful in clinical research and clinical practice. As to the former, speech-language pathologists are specialists in behavioral modification, and many of the symptoms included in the SWAL–QOL should be responsive to effective behavioral modification. As to the latter, a standardized symptom battery could aid the busy clinician in ascertaining specific problems patients are experiencing as well as monitoring symptom severity over time. The three symptom severity scales may prove to be clinically useful ways to analyze symptom severity at the group level.

Results for clinical validity mesh very well with those for factorial validity and suggest a hierarchy of dysphagia-specific quality of life. Across the results of clinical validity, the scales tapping food selection, burden, mental health, and social functioning were most sensitive to differences in the severity of dysphagia as clinically defined. These same scales were also the primary contributors to dysphagia-specific quality of life as derived from the factor analysis (see Table 8). This synergy across different validity tests suggests that, as more data are collected on the SWAL–QOL and the SWAL–CARE, we may be able to offer measurement priorities for efficacy and effectiveness studies if investigators are unable to implement the 44-item SWAL–QOL and the 15-item SWAL–CARE. For the time being, however, we offer the SWAL–QOL as ten distinct scales and the SWAL–CARE as three distinct scales. We appreciate that some clinicians and investigators desire a single score for outcome assessment. However, we feel it is premature at this time to derive a total score across the different scales. Further, interpretation problems have been noted with summary scores derived from other quality-of-life tools [84–86]. As more data are gathered on the SWAL–QOL and the SWAL–CARE, we hope to offer a more parsimonious alternative to the current profile of scales.

There are limitations of our work that should be noted. First, our sample of dysphagic patients was not severely impaired, as defined by the food texture of their meals, the consistency of their liquid intake, and tube-feeding status. However, they did report many symptoms of dysphagia (see Table 16), and their quality-of-life scores were considerably diminished compared with the normal control group (see Table 12). While we sampled patients across six national sites, our sample was largely white and male, the latter reflective of sampling patients at VA medical centers. We did conduct sensitivity tests and did not find the integrity of the scales (item convergent and discriminant validity and scale reliability) to differ by gender or race (available upon request). However, we do caution that the data reported here should not be considered as any type of national norm because of the over-representation of men and Caucasians.

Palliation of symptoms, and thus maintenance of quality of life, is an important goal of clinical care for dysphagic patients. As Siegrist and Junge [87] argue: "Quality of life data promote more humane, personal attitudes toward patients." We developed the SWAL–QOL and the SWAL–CARE to enhance the measurement armamentarium of dysphagia researchers and clinicians, and thus their research and clinical ingenuity and skills. It is intended that the standardization and publication of the SWAL–QOL and the SWAL–CARE will facilitate their use in clinical research and clinical practice to better understand treatment effectiveness as a critical step toward improving patients' quality of life and quality of care.

Acknowledgments. This work was supported by the Department of Veterans Affairs (HSR&D HFP #96-001; RR&D C-2016; IIR #95-033 to Dr. McHorney). The authors gratefully acknowledge Jody McIntyre for her innumerable contributions to the SWAL–QOL project. The authors also acknowledge the contributions of Jeri Logemann, Colleen Clarke, Bonnie Martin–Harris, Carol Blossom Stach, Steven Deutsch, and Thomas Sowell for data collection for the field test.

References

1. Katz S, Ford AB, Moskowitz RW, Jacobson BA, Jaffe MW: Studies of illness in the aged: The index of ADL: A standardized measure of biological and psychosocial function. *JAMA 185:*914–919, 1963
2. Katz S, Akpom CA: A measure of primary sociobiological functions. *Int J Health Serv 6:*493–508, 1976
3. Sonn U: Longitudinal studies of dependence in daily life activities among elderly persons. *Scand J Rehabil Med 34*(Suppl)*:*1–35, 1996
4. Dunlop D, Hughes S, Manheim L: Disability in activities of daily living: Patterns of change and a hierarchy of disability. *Am J Public Health 87:*378–383, 1997
5. Sonn U, Svenson E: Measures of individual and group changes in ordered categorical data: Application to the ADL staircase. *Scand JRehabil Med 29:*233–242, 1997
6. Berberger–Gateau P, Rainville C, Letenneur L, Dartigues J: A hierarchical model of domains of disablement in the elderly: A longitudinal approach. *Disabil Rehabil 22:*308–317, 2000
7. Kempen G, Suurmeijer T: The development of a hierarchical polychotomous ADL–IADL scale for noninstitutionalized elders. *Gerontologist 30:*497–502, 1990
8. Suurmeifer P, Kempen G: Behavioural changes as an outcome of disease: The development of an instrument. *Int J Health Sci 1:*189–194, 1990
9. Crimmins EM, Saito Y: Getting better and getting worse. *J Aging Health 5:*3–36, 1993
10. Crimmins E, Hayward M, Saito Y: Differentials in active life expectancy in the older population of the United States. *J Gerontol 51B:*S111–S120, 1996
11. Kempen G, Miedema I, Ormel J, Molenaar W: The assessment of disability with the Groningen Activity Restriction Scale. Conceptual framework and psychometric properties. *Soc Sci Med 43:*1601–1610, 1996
12. Wolinsky F, Stump T, Callahan C, Johnson R: Consistency and change in functional status among older adults over time. *J Aging Health 8:*155–182, 1996
13. Ferrucci L, Guralnik J, Cecchi F, Marchionni N, Salani B, Kasper J, Celli R, Giardini S, Heikkinen E, Jylha J, Baroni A: Constant hierarchic patterns of physical functioning across seven populations in five countries. *Gerontologist 38:*286–294, 1998
14. Spector W, Fleishman J: Combining activities of daily living with instrumental activities of daily living to measure functional disability. *J Gerontol 53:*S46–S57, 1998
15. Thomas V, Rockwood K, McDowell I: Multidimensionality in instrumental and basic activities of daily living. *J Clin Epidemiol 51:*315–321, 1998
16. Robbins J, Levine R, Maser A, Rosenbek J, Kempster GB: Swallowing after unilateral stroke of the cerebral cortex. *Arch Phys Med Rehabil 74:*1295–1300, 1993
17. Bell RC, Low LH, Jackson HJ, Dudgeon PL, Copolov DL: Latent trait modelling of symptoms of schizophrenia. *Psychol Med 24:*335–345, 1994
18. Rosenbek J, Robbins J, Roecker E, Coyle M, Wood J: A penetration–aspiration scale. *Dysphagia 11:*93–98, 1996
19. Gustafsson B, Tibbling L: Dysphagia: An unrecognized handicap. *Dysphagia 6:*193–199, 1991
20. McHorney C, Rosenbek J: Functional outcomes assessment of adults with oropharyngeal dysphagia. *Semin Speech Lang 19:*235–247, 1998
21. Langmore S: An important tool for measuring quality of life. *Dysphagia 15:*134–135, 2000
22. McHorney C, Bricker D, Kramer A, Rosenbek J, Robbins J, Chignell K, Logemann J, Clarke C: The SWAL–QOL outcomes tool for oropharyngeal dysphagia in adults: I. Conceptual foundation and item development. *Dysphagia 15:*115–121, 2000
23. McHorney C, Bricker D, Robbins J, Kramer A, Rosenbek J, Chignell K: The SWAL–QOL outcomes tool for oropharyngeal dysphagia in adults: II. Item reduction and preliminary scaling. *Dysphagia 15:*122–133, 2000
24. Good–Fratturelli M, Curlee R, Holle J: Prevalence and nature of dysphagia in VA patients with COPD referred for videofluoroscopic swallow examination. *J Commun Disord 33:*93–110, 2000
25. Nicosia M, Hind J, Roecker E, Carnes M, Doyle J, Dengel G, Fagerholm M, Robbins J: Age effects on temporal evo-

lution of isometric and swallowing pressure. *J Gerontol 55A:*M634–M640, 2000

26. Heberlein TA, Baumgartner R: Factors affecting response rates to mailed questionnaires: A quantitative analysis of the published literature. *Am Sociol Rev 43:*447–462, 1978
27. Nunnally JC: *Psychometric Theory*. New York: McGraw–Hill, 1978
28. Manning WG, Newhouse JP, Ware JE: The status of health in demand estimation: Beyond excellent, good, fair, and poor. In: Fuchs VR (ed.): *Economic Aspects of Health*. Chicago: University of Chicago Press, 1982
29. McHorney CA, Ware JE, Rogers W, Raczek A, Lu JFR: The validity and relative precision of MOS short- and long-form health status scales and Dartmouth COOP Charts: Results from the Medical Outcomes Study. *Med Care 30:*MS253–MS265, 1992
30. McHorney C: Generic health measurement: Past accomplishments and a measurement paradigm for the 21st century. *Ann Intern Med 127:*743–750, 1997
31. Sprangers M, Aaronson N: The role of health care providers and significant others in evaluating the quality of life of patients with chronic disease: A review. *J Clin Epidemiol 45:*743–760, 1992
32. Likert R: A technique for the measurement of attitudes. *Arch Psychol 140:*5–55, 1932
33. Comrey AL: *A First Course in Factor Analysis*. New York: Academic Press, 1973
34. Stewart AL, Ware JE: *Measuring Functioning and Well-Being: The Medical Outcomes Study Approach*. Durham, NC: Duke University Press, 1992
35. Hays RD, Hayashi T: Beyond internal consistency reliability: Rationale and user's guide for multitrait analysis program on the microcomputer. *Behav Res Meth, Instrum Comput 22:*167–175, 1990
36. Samejima F: Estimation of latent ability using a response pattern of graded scores. *Psychometr Monogr 17:*1–100, 1969
37. Samejima F: Graded response model. In: van der Linden W, Hambleton R (eds.): *Handbook of Modern Item Response Theory*, New York: Springer, 1997
38. McHorney C, Tarlov A: The use of health status measures for individual patient level applications: Problems and prospects. *Qual Life Res 3:*43–44, 1995
39. Cronbach LJ: Coefficient alpha and the internal structure of tests. *Psychometrika 16:*297–334, 1951
40. DeVellis R: *Scale Development: Theory and Applications*. Newbury Park, CA: SAGE Publications, 1991
41. Fleiss JL: *Statistical Methods for Rates and Proportions*. New York: John Wiley & Sons, 1981
42. Stewart AL, Hays RD, Ware JE: The MOS Short-Form General Health Survey: Reliability and validity in a patient population. *Med Care 26:*724–735, 1988
43. Ware JE, Sherbourne CD: The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection. *Med Care 30:*473–483, 1992
44. Hall JA, Feldstein M, Fretwell MD, Rowe JW, Epstein AM: Older patients' health status and satisfaction with medical care in an HMO population. *Med Care 28:*261–270, 1990
45. Kane R, Maciejewski M, Finch M: The relationship of patient satisfaction with care and clinical outcomes. *Med Care 35:*714–730, 1997
46. Wensing M, Grol R, Asberg J, van Montfort P, van Weel C, Felling A: Does the health status of chronically ill patients predict their judgments of the quality of general practice care? *Qual Life Res 6:*293–299, 1997
47. Hall J, Roter D, Milburn M, Daltroy L: Why are sicker patients less satisfied with their medical care? Tests of two explanatory models. *Health Psychol 17:*70–75, 1998
48. Kerlinger FN: *Foundations of Behavioral Research*. New York: Holt, Rinehart and Winston, 1973
49. Hattie J, Cooksey R: Procedures for assessing the validities of tests using the known-groups method. *Appl Psychol Meas 8:*295–305, 1984
50. McHorney C: Methodological and psychometric issues in health status assessment across populations and applications. In: Albrecht GL, Fitzpatrick R (eds.) *Advances in Medical Sociology*. Stamford, CT: JAI Press, Inc., 1994
51. Landis JR, Koch GG: The measurement of observer agreement for categorical data. *Biometrics 33:*159–174, 1977
52. Berwick DM, Murphy JM, Goldman PA, Ware JE, Barsky AJ, Weinstein MC: Performance of a five-item mental health screening test. *Med Care 29:*169–176, 1991
53. Holmes W: A short, psychiatric, case-finding measure for HIV seropositive outpatients—performance characteristics of the 5-item mental health subscale of the SF-20 in a male, seropositive sample. *Med Care 36:*247–243, 1998
54. Logemann J: *Evaluation and Treatment of Swallowing Disorders*. Austin, TX: Pro-Ed, 1998
55. Wainer H, Braun HI: *Test Validity*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1988
56. Hind JA, Robbins JA: Lingual exercise protocol for head and neck cancer: A case study. Ninth Annual Meeting of the Dysphagia Research Society, Savannah, GA, 2000
57. Rubin H, Gandek B, Rogers WH, Kosinski M, McHorney C, Ware JE: Patients' ratings of outpatient visits in different practice settings: Results from the Medical Outcomes Study. *JAMA 270:*835–840, 1993
58. Wensing M, Grol R, Smits A: Quality judgments by patients on general practice care: A literature analysis. *Soc Sci Med 38:*45–53, 1994
59. Thomas L, McColl E, Priest J, Bond S, Boys R: Newcastle satisfaction with nursing scales: An instrument for quality assessments of nursing care. *Qual Health Care 5:*67–72, 1996
60. Elliott–Burke T, Pothast L: Measuring patient satisfaction in an outpatient orthopedic setting, Part 1: Key drivers and results. *J Rehabil Outcomes Meas 1:*18–25, 1997
61. Batalden P, Mohr J, Nelson E, Plume S: Improving health care, Part 4: Concepts for improving any clinical process. *Jt Comm J Qual Improv 22:*651–659, 1996
62. Nelson E, Mohr J, Batalden P, Plume S: Improving health care, Part 1: The clinical value compass. *Jt Comm J Qual Improv 22:*243–258, 1996
63. Pascoe G: Patient satisfaction in primary health care: A literature review and analysis. *Eval Program Plann 6:*185–210, 1983
64. Sherbourne CD, Hays RD, Ordway L, Dimatteo MR, Kravitz RL: Antecedents of adherence to medical recommendations: Results from the Medical Outcomes Study. *J Behav Med 15:*447–469, 1992
65. Harris L, Luft F, Rudy D, Tierney W: Correlates of health care satisfaction in inner-city patients with hypertension and chronic renal insufficiency. *Soc Sci Med 41:*1639–1645, 1999
66. Marquis MS, Davies AR, Ware JE: Patient satisfaction and change in medical care provider: A longitudinal study. *Med Care 21:*821–829, 1983

67. Ware JE, Davies AR: Behavioral consequences of consumer dissatisfaction with medical care. *Eval Program Plann 6:*291–297, 1983
68. Sorensen A, Wersinger R: Factors influencing disenrollment from an HMO. *Med Care 19:*766–773, 1981
69. Hennelly V, Boxerman S: Disenrollment from a prepaid group plan: A multivariate analysis. *Med Care 21:*1154–1167, 1983
70. Grazier KL, Richardson WC, Martin DP, Diehr P: Factors affecting choice of health care plans. *Health Serv Res 20:*659–682, 1986
71. Harrington C, Newcomer R, Preston S: A comparison of S/HMO disenrollees and continuing members. *Inquiry 30:*429–440, 1993
72. Newcomer R, Preston S, Harrington C: Health plan satisfaction and risk of disenrollment among Social/HMO and fee-for-service recipients. *Inquiry 3:*144–154, 1996
73. Etter J, Perneger T: Validating a satisfaction questionnaire using multiple approaches: A case study. *Soc Sci Med 45:*879–885, 1997
74. Hennelly V, Boxerman S: Out-of-plan use and disenrollment: Outgrowths of dissatisfaction with a prepaid group plan. *Med Care 21:*348–359, 1983
75. Senf J, Weiss B: Patient satisfaction with health care: Intentions and change in plan. *Eval Program Plann 14:*299–306, 1991
76. Kerr E, Hays R, Lee M, Siu A: Does dissatisfaction with access to specialists affect the desire to leave a managed care plan? *Med Care Res Rev 55:*59–77, 1998
77. Wells KB, Hays RD, Burnam MA, Rogers W, Greenfield S, Ware JE: Detection of depressive disorder for patients receiving prepaid or fee-for-service care: Results from the Medical Outcomes Study. *JAMA 262:*3298–3302, 1989
78. Potts MK, Burnam MA, Wells KB: Gender differences in depression detection: A comparison of clinician diagnosis and standardized assessment. *Psychol Assess 3:*609–615, 1991
79. Kroenke K: Discovering depression in medical patients: Reasonable expectations. *Ann Intern Med 126:*463–465, 1997
80. Callahan C, Hui S, Nienaber N, Musick B, Tierney W: Longitudinal study of depression and health services use among elderly primary care patients. *J Am Geriatr Soc 42:*833–838, 1994
81. Luber M, Hollenberg J, Williams–Russo P, DiDomenico T, Myers B, Alexopoulos G, Charlson M: Diagnosis, treatment, comorbidity, and resource utilization of depressed patients in general medical practice. *Int J Psychiatry Med 30:*1–13, 2000
82. Wells KB, Stewart A, Hays RD, Burnam MA, Rogers W, Daniels M, Berry S, Greenfield S, Ware JE: The functioning and well-being of depressed patients: Results from the Medical Outcomes Study. *JAMA 262:*914–919, 1989
83. Idler E, Benyamini Y: Self-rated health and mortality: A review of twenty-seven community studies. *J Health Soc Behav 38:*21–37, 1997
84. Jenkinson C, Phil D, Gray A, Doll H, Lawrence K, Keoghane S, Layte R: Evaluation of index and profile measures of health status in a randomized controlled trial. *Med Care 35:*1109–1118, 1997
85. McHorney C: Methodological inquiries in health status assessment. *Med Care 36:*445–448, 1998
86. Simon G, Revicki D, Grothaus L, Vonkorff M: SF-36 summary scores. Are physical and mental health truly distinct? *Med Care 36:*567–572, 1998
87. Siegrist J, Junge A: Background material for the workshop on QALYS. *Soc Sci Med 29:*463–468, 1989