REVIEW

A review of functional status measures for workers with upper extremity disorders

D F Salerno, C Copley-Merriman, T N Taylor, J Shinogle, R M Schulz

....

In order to identify functional status measures for epidemiological studies among workers with mild to moderate disorders of the neck and upper extremity, a literature search was conducted for the years 1966 to 2001. Inclusion criteria were: (1) relevance to neck and upper extremity; (2) assessment among workers; and (3) relevance to mild to moderate disorders. Of 13 instruments reviewed, six measures were tested among workers. The three best measures, depending on the purpose of research, included the standardised Nordic Musculoskeletal Questionnaire, the Upper Extremity Questionnaire, and the Neck and Upper Limb Instrument. Development of a functional protocol is regarded as a realistic enhancement for research of neck and upper extremity disorders in the workplace. For research and clinical practice, measures of functional status, sensitive enough to measure the subtle conditions in mild to moderate disorders, may provide prognostic information about the risk of developing musculoskeletal disorders in apparently healthy patients. Appropriate use of functional status questionnaires is imperative for a meaningful portrayal of health.

....

F unctional status measures can correlate pain to performance, with direct relevance to employers and workers. Although functional status may range from full ability to severe disability, few measures have been designed for relatively healthy, active workers with upper extremity disorders,[1] particularly those with mild to moderate conditions.

Over 75 functional status instruments exist for patients with disability from arthritis or diabetes,[2] yet most have focused on severe disability. Thus, although useful in certain epidemiological investigations, most measures are insufficient in addressing the problem of mild to moderate conditions. However, research shows that workers without discernible medical diagnosis often report interference with activities at work or home.[3] Investigators need measures capable of detecting subtle as well as pronounced musculoskeletal conditions, and the impact of these conditions on performance.

Musculoskeletal disorders are believed to represent the largest category of work related illness in Britain.[4] In the United States (private sector), nearly 6 million workers experience non-fatal injuries or illnesses.[5] Although musculoskeletal

disorders are among the most prevalent and symptomatic complaints among workers, occupational medicine lacks measures for certain disorders, particularly in early stages.[6]

New research protocols typically include not only clinical laboratory tests, but also self reported questionnaires.[7–9] This review identifies functional status instruments easily used in occupational health surveys among a working population of mainly healthy subjects. It identifies measures for workers with mild to moderate disorders of the neck and upper extremity. The review is limited to measures for employed adults (distinct from workers who are not employed, or workers under 18 years of age).

DEFINING FUNCTIONAL STATUS

Functional status has been characterised as health status, activities of daily living, level of impairment, disability, or handicap. In general, the construct of "function" contains physical, emotional, and social attributes.[10] For this review, a functional status measure is defined as an instrument to assess how health and strength, vitality, symptoms (for example, pain or discomfort), emotion, or desires affect performance of everyday activities, recreation, social relations, and work. In short, how physical conditions affect activity.

Stock and colleagues[1] identified 12 functional domains relevant to workers: work, household and family responsibilities, self care, transportation, sexual activity, sleep, social activities, recreational activities, mood, self esteem, financial effects, and the iatrogenic effects of assessment and treatment.

CONCEPTUAL FRAMEWORK FOR FUNCTIONAL STATUS MEASURES

Functional status measures may be generic (general health) or specific (disease related); discriminative (determining if the condition is better or worse) or evaluative (measuring whether a score has changed).[11] Various classification schemas have been proposed,[12][13] yet there is no consensus

See end of article for authors' affiliations

....

Correspondence to:
Dr D F Salerno, 2800
Plymouth Road, Ann Arbor,
MI 48105, USA;
Deborah.Salerno@
pfizer.com

Accepted 12 March 2002
....

....

Abbreviations: ADL, activities of daily living; CTS, carpal tunnel syndrome; DASH, Disabilities of the Arm, Shoulder, and Hand; DRI, Disability Rating Index; FSS, Functional Status Scale; ICC, intraclass correlation; MFA, Musculoskeletal Functional Assessment; MHQ, Michigan Hand Outcomes Questionnaire; NDI, Neck Disability Index; NMQ, Nordic Musculoskeletal Questionnaires; NULI, Neck and Upper Limb Instrument; SPADI, Shoulder Pain and Disability Index; SRM, standardised response mean; SSS, Symptom Severity Scale; UEQ, Upper Extremity Questionnaire; VAS, visual analogue scale

Main messages

- Functional status measures can correlate pain to performance.
- Few self reported measures have been designed specifically for workers.
- The impact of mild to moderate disorders on the workforce is unknown.
- Three measures are identified for epidemiological studies among workers.
- Consistent use of functional status measures is encouraged.

Policy implications

- Uniform data collection provides the benefit of a standard reporting environment.
- Global data standards enable consistently high quality reports, for intelligible comparisons across industries.

on the basis for stratification of measures. This review focuses on self reported measures, in contrast to clinician based or economic measures (fig 1). It is important to distinguish between instruments classifying diagnosis or pain, and those classifying functional status.

Ideally, functional measures would provide data on a full spectrum of function. An important challenge is to measure subtle disease entities that are troublesome to workers (and employers) but more difficult to evaluate. This new approach requires quantification of symptoms below the threshold of those traditionally measured via clinical laboratory tests or physical examination.

METHODS

To identify self reported functional status instruments for neck and upper extremity disorders among workers, a Medline search was conducted for English language publications between the years 1966 and 2001. Keywords included:

carpal tunnel syndrome, functional status, health surveys, musculoskeletal, occupational health, outcome measures, questionnaire, neck, upper extremity, and worker.

The following criteria were used to select self reported instruments for review: (1) relevance to neck and upper extremity conditions (indicated by question content); (2) assessment among workers; and (3) relevance to mild to moderate disorders (mild to moderate conditions could be defined by the patient, not necessarily correlated with abnormal laboratory tests or physical examination). Instruments were selected for this review if published in peer reviewed studies explicitly designed to evaluate psychometric properties of validity, repeatability, and responsiveness to change.

Psychometric properties
Psychometric properties were defined as follows.

Validity
How well an instrument measures what it is supposed to measure, how it reflects reality. A valid scale provides for accurate inferences.[14] Validity is often quantified by correlation analyses, ROC curve calculations, regression models, or estimates of sensitivity and specificity. Construct validity shows how the new measure compares with other associated measures. Criterion validity comparisons are with some gold standard.[15] Content or face validity is the extent to which a set of items reflects a certain domain. Evidence of validity is found by repeated use of the instrument with performance as expected.

Repeatability
Consistency over time (test-retest, inter- and intrarater reliability). For dichotomous data, the odds ratio or kappa statistics are recommended. Kappa statistics account for agreement by chance.[16] For ordered data, weighted kappa values are used.[17] A kappa value less than 0.40 represents poor agreement beyond chance, between 0.40 and 0.75 is fair to good; and a value greater than 0.75 is excellent.[18] For continuous measures, the intraclass correlation (ICC) is the statistical
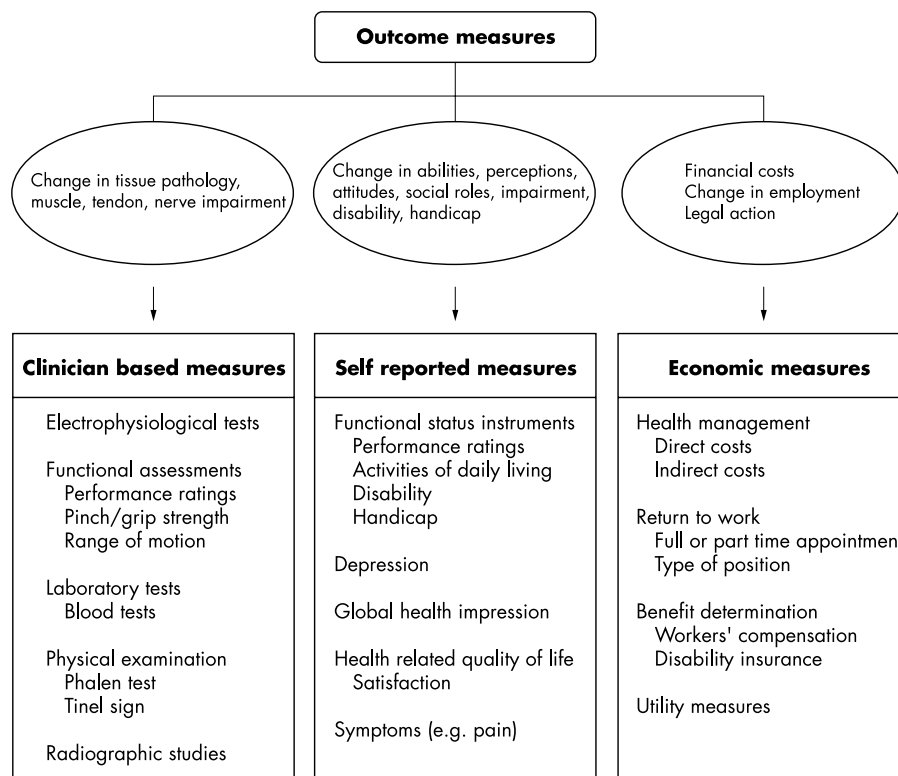

Figure 1 Framework for outcome measures.

analogue, combining a measure of correlation with a test in the difference of means.[19] Although Pearson product-moment correlations can be used, it is known that observations may disagree sharply yet still be correlated.[20] Measures may be compared between instruments only if they have been calculated from comparable populations.

Internal consistency, the ability to measure a single concept, is measured by Cronbach alpha.[21] A value of 0.70 is good, 0.80 is sufficient, and 0.90 is excellent.[22]

Responsiveness to change

The ability to detect change over time. Responsiveness is commonly quantified by effect size, or the standardised response mean (SRM). Cohen[23] defined the effect size statistic, d, as the difference between means divided by the standard deviation of either group. An effect size of 0.20 or less is small, 0.50 is moderate, and a value of 0.80 or greater is large. The larger the effect size, the more responsive the instrument.

The SRM is the mean change in scores from baseline to follow up divided by the standard deviation of changes. Interpretation of values is similar to effect size. When the correlation between baseline and follow up scores is equal to 0.5, the SRM is equal to the effect size. When the correlation is high, the SRM is greater than the effect size; when the correlation is low, the effect size is as much as 1.4 times higher than the SRM.[24] Again, responsiveness depends on having similar populations being assessed between instruments.

RESULTS

A total of 13 self reported neck and upper extremity instruments were reviewed. Based on appearance, all instruments had content validity. Seven instruments were tested among surgical patients. Six instruments had been tested among workers, including three instruments relevant for workers with mild to moderate disorders.

Functional instruments tested among surgical patients

Shoulder Pain and Disability Index (SPADI)[25]

This has 13 items on 10 cm visual analogue scales (VAS). Ratings include pain severity, difficulty carrying heavy objects, and placing things on a high shelf. Among 37 men, there was moderate to high correlation (Pearson r range: −0.55 to −0.80) with shoulder range of motion. Among 23 subjects in the test-retest study, ICC was 0.66. In assessing responsiveness, the group showed a mean decrease (−25.6) in scores (possible range: 0–100).

Shoulder Rating Questionnaire[26]

This has 21 items including a global assessment on a 10 cm VAS, questions regarding pain, daily activities, recreational activities, work, satisfaction, and area for improvement graded on five point scales. In a study among 100 patients, there was moderate to high correlation to the revised Arthritis Impact Measurement Scales.[27] There were significant correlations for global assessment (Spearman r = −0.56), level of satisfaction (r = −0.56), daily activities (r = −0.84), pain (r = −0.86), and work (r = −0.89). Cronbach alpha ranged from 0.71 to 0.90; weighted kappa were greater than 0.70. Tests for responsiveness (n = 30) showed standardised response means (SRM) between 1.1 and 1.9.

Symptom Severity Scale (SSS)[28] [29]

This has 11 items on five point scales assessing pain severity, nocturnal occurrence of pain, frequency, duration, numbness, weakness, tingling, difficulty with grasping, and use of small objects. Testing was conducted in a three month prospective study, and a retrospective study of patients evaluated after surgery. There was significant correlation for pinch strength (Spearman r = 0.47) and grip strength (r = 0.38). Test-retest reliability was high (Pearson r = 0.91) (n = 31), as was internal consistency (Cronbach α = 0.89) (n = 67). The effect size for SSS was 1.13 in the prospective cohort (n = 26), and 1.4 in the retrospective cohort (n = 38).

Functional Status Scale (FSS)[28]

This has items assessing eight activities (writing, buttoning clothes, holding a book, gripping a telephone, opening a jar, household chores, carrying grocery bags, bathing, and dressing) on a fivepoint scale. Psychometric testing was conducted in a prospective study, and a retrospective study of patients after surgery for carpal tunnel syndrome (CTS). The highest correlations were for pinch strength (Spearman r = 0.60), and grip strength (r = 0.50). Test-retest reliability was high (Pearson r = 0.93) (n = 31) as was internal consistency (Cronbach α = 0.91) (n = 67). The effect size was 0.71 in the prospective cohort (n = 26), and 0.82 in the retrospective group (n = 38).

Disabilities of the Arm, Shoulder and Hand (DASH)[30–32]

This has 30 items assessing symptoms: daily activities, recreation, self care, sleep, sports, family care, occupation, socialising, and self image. It includes dichotomous items, and five to six point scales. The DASH correlated highly (>−0.75) with other measures of function, disability, and pain. Both test-retest reliability and internal consistency exceeded 0.95. Optional modules are available for athletes/performing artists, and working populations. The DASH was developed by organisations for surgeons to measure disability and symptoms.

Michigan Hand Outcomes Questionnaire (MHQ)[33] [34]

This has 37 items on six scales: overall hand function, activities of daily living (ADLs), pain, work performance, aesthetics, and patient satisfaction after surgery. In comparing three MHQ scales to the SF-12, a moderate correlation was found (range 0.54–0.79). Tests of validity showed a significant difference for patients with CTS versus arthritis. Spearman correlation ranged from 0.81 to 0.97. Cronbach α ranged from 0.86 to 0.97.

Short Musculoskeletal Functional Assessment Questionnaire[35]

This is based on the Musculoskeletal Function Assessment (MFA).[36] [37] It has 46 items in two indices: a dysfunction index and bother index (how much patients are bothered by problems) with five point scales. Among 420 patients, both indices had significant correlations with walking speed and grip strength (Pearson $r \geqslant 0.40$), ADLs, recreational activities, and emotional function ($r \geqslant 0.40$), and with the SF-36 subscales. Test-retest reliability was high (ICC = 0.93, dysfunction index; ICC = 0.88 bother index), as was internal consistency (Cronbach α > 0.92). A test for responsiveness showed SRMs between 0.76 and −1.14.

Functional instruments tested among workers

Six questionnaires were designed for and tested among workers, of which three measures were relevant for mild to moderate disorders: the standardised Nordic Musculoskeletal Questionnaire, the Upper Extremity Questionnaire, and the Neck and Upper Limb Instrument (table 1).

Neck Disability Index (NDI)[38]

This was tested among 48 subjects with neck pain in a chiropractic clinic. A subsequent study[39] involved a larger sample of 237 neck pain patients, and showed similarly high internal consistency (Cronbach α = 0.92). Exploratory factor analysis

Table 1 Description of selected functional status instruments tested among workers

	Description	Validity	Reliability	Responsiveness to change
1	**Neck Disability Index** (NDI): 10 items Pain intensity, personal care, lifting, reading, headaches, concentration, work, driving, sleeping, recreation. Response: 1 (poorest) to 5 (best) [Vernon and Mior, 1991]	Peer review, patient feedback. Correlation of 0.6 between changes in scores pre- and post-treatment compared to VAS for activity levels. Correlation of 0.7 between NDI and McGill Pain Questionnaire.	Pearson correlation: 0.89 Cronbach alpha: 0.80 total index Cronbach alpha: each item above 0.76	33% change in pre- v post-treatment scores
2	**Disability Rating Index** (DRI): 12 items Dressing, outdoor walks, climbing stairs, sitting long time, standing over sink, carrying bag, making bed, running, light work, heavy work, lifting heavy objects, exercise/sports. Response: 10 cm VAS: without difficulty – not at all [Salén et al, 1994]	Construct validity: median scores very low (0.9% and 8.8%) in healthy groups, high (50–65%) among patients with arthritis. Pearson intraclass correlation to observers' assessment of performance=0.48.	Test-retest reliability 0.92 for 3 day interval 0.95 for 1 day interval Inter- 0.99 and intra-rater: 0.98 Cronbach alpha 0.84	16% change in median pre- v postoperative scores
3	**Upper Extremity Function Scale**: 8 items Sleeping, writing, opening jars, picking up small objects with fingers, driving car >30 min, opening door, carrying milk jug from refrigerator, washing dishes. Response: 1 (no problem) to 10 (major problem) [Pransky et al, 1997]	Able to discriminate between levels of severity (significant t test value) for working status, duration of symptoms, effect on job ability, Phalen's test. Correlation with average pain level r=0.67, fear of pain, r=0.44, pinch strength r=−0.40	Cronbach alpha: 0.83 to 0.93	Upper extremity cohort SRM range: −1.33 to −0.53 CTS cohort SRM range: −1.03 to 0.92
4	**Nordic Musculoskeletal Questionnaires**: 27 items 3 questions: Have you had trouble (ache, pain, discomfort, numbness), Had trouble during last seven days; Been prevented from normal activities in 9 body areas (neck, shoulder, elbows, wrists/hands, upper back, lower back, hips/thighs/buttocks, knees, ankles/feet). Response: Yes/No, multiple choice [Kuorinka et al, 1987]	Tests with clinical history: non-identical answers varied between 0–20%. Comparison to PT assessment: percentage of disagreement varied between 0–13%.	*Neck-shoulder questionnaire* disagreeing responses varied from 0–30% *Low back questionnaire* disagreeing responses varied from 0–4% (1 item had 25% variance and reformulated in the final version)	N/A
5	**Upper Extremity Questionnaire**: 123 items Medical history, exercise, symptoms (burning, stiffness, pain, cramping, tightness, aching, soreness, tingling, numbness) in 3 body areas (neck/shoulder/upper arm, elbow/forearm, and wrist/hand/finger); frequency, severity, interference with work. Response: Yes/No, 10 cm VAS, multiple choice [Franzblau et al, 1997; Salerno et al, 2001]	N/A	Kappa: 0.56–1 medical history Symptom kappa values 0.39–0.86 neck/shoulder/upper arm 0.44–0.86 elbow/forearm 0.44–0.84 wrist/hand/fingers	N/A
6	**Neck and Upper Limb Index** (NULI): 20 items 7 physical activity items, 4 work-related, 6 psychosocial, 2 sleep related, 1 item related to negative effects of evaluation or treatment. Response: 7 pt scales [Stock et al, 2000]	Correlation to Sickness Impact Profile, question on overall functional status, PT perception of functional status, hand grip strength, shoulder abduction, scratch test.	Test-retest: ICC=0.88 Ontario cohort; ICC=0.83 Quebec cohort. Cronbach alpha: 0.90 Ontario Cronbach alpha: 0.93 Quebec	SRM: 1.48 Ontario cohort SRM: 1.63 Quebec cohort

VAS, visual analogue scale; SRM, standardised response mean; CTS, carpal tunnel syndrome; PT, physical therapist; N/A, not available.

indicated high loadings on items of work, driving, and recreational activities, measuring more of the physical aspect of pain disability.

Disability Rating Index (DRI)[40]
This was tested among 1092 healthy blue and white collar workers, and 366 patients with pain in the neck, shoulder, or low back with different levels of ability. Responsiveness was significant among 19 arthritis patients with median preoperative scores of 52%, and median postoperative scores of 36%.

Upper Extremity Function Scale[41]
This was tested among 108 patients with upper extremity disorders receiving Workers' Compensation, and 165 patients at a hand clinic. The instrument correlated highly with average pain level (Pearson $r = 0.67$) and fear of pain ($r = 0.44$). Among the CTS patients, it correlated highest with worst pain level ($r = 0.54$), and pinch strength ($r = -0.40$). Cronbach α ranged from 0.83 to 0.93 across study groups. The SRM ranged from −0.53 to −1.33 for a subgroup of 16 patients who reported being significantly better; SRM ranged from 0.39 to −1.03 for a subgroup of 55 CTS patients.

Nordic Musculoskeletal Questionnaires (NMQ)[42 43]
These have been tested among 27 clerical workers, 82 women in electronics manufacturing, 17 medical secretaries, 22 railway maintenance workers, and 29 safety engineers. A specific Neck and Shoulder Questionnaire assesses the severity of symptoms in terms of their effect on activities at work and during leisure time. Widely used in Europe, the NMQ was adapted by the National Institute for Occupational Safety and Health (NIOSH) in the United States.[44]

Upper Extremity Questionnaire (UEQ)[45 46]
This has been tested for test-retest reliability among 148 manufacturing workers, and 138 keyboard operators in the United States. For symptom reports among keyboard operators, most kappa values were between 0.60 and 0.89.[47] Symptom severity and interference with production rates and/or usual standard of quality were less stable. Among the psychosocial measures, Perceived Stress and Job Dissatisfaction Scales were most consistent (ICC = 0.88); coworker support was least consistent (ICC = 0.44), perhaps due to workforce characteristics, such as high turnover.

Neck and Upper Limb Instrument (NULI)[48]
This was designed for research or clinical use among workers. In a test-retest assessment among 99 subjects, ICC was 0.88. The tool is used to evaluate effectiveness of interventions and prevention of disability.[49]

DISCUSSION
This review identified 13 self reported functional status measures, and the context for their use. While the measures have been used, not all have been *tested* among workers. None of the seven measures designed for patients undergoing surgery was tested among patients receiving conservative treatment. Of the six measures tested among workers, three measures were relevant for mild to moderate neck and upper extremity conditions: the NMQ, Upper Extremity Questionnaire, and the NULI.

Considerations for a worker population
Basic psychometric principles demand that measures be tested in the population in which they will be used. Implicit in the objective to find measures for workers is the assumption that functional status among workers is different to that in a clinical population. Floor and ceiling effects (that is, the inability of an instrument to accurately reflect patients at the ends of the spectrum) should be considered since questionnaires designed for clinical populations may not adequately address the more subtle conditions of active workers.

Age or gender distributions in the workplace can make statistical comparisons difficult if the sample size is small.[50] For example, traditional gender concentrations in manufacturing and service industries pose statistical problems for generalisability.

In addition, psychosocial and vocational factors need to be considered,[51 52] as factors such as job dissatisfaction, labour relations, dependents, or layoffs may affect outcomes. Selection issues arise if workers believe health status could affect their ability to receive benefits (for example, public or private disability insurance, or health insurance).[53]

Another concern is proper comparison of normative data among workers. Most normative values are based on convenience samples with certain limitations, namely spectrum bias.[54] For example, nerve conduction studies among workers are routinely compared with normative data. Research shows that norms among workers are different from conventional clinical norms, and need attention to avoid misclassification.[55]

One reason that instruments have rarely been tested among active workers is the logistical difficulty of arranging access. Measures for mild to moderate conditions are a further challenge, as small changes are harder to detect. However, these conditions are gaining importance with the growth of industries that rely on static or constrained postures of the neck, and/or repetitive use of the upper extremity.

Implications for research and clinical trials
Overall, intervention for subtle conditions may prove cost effective and prevent escalating morbidity. Although minor physical damage is reversible, a cascade of biochemical and mechanical changes falls into place in response to injury.[56] These changes lead to immune responses and inflammation, which can be a precursor to disability. In early stages, medical care, physical therapy, massage therapy, or pharmacological interventions may provide proper health management. Obviously, the advantage of detecting mild to moderate disorders is that conservative interventions (for example, engineering or administrative controls) may be initiated to minimise decrements in function.

Although there has been progress in standardising a core set of classification criteria in upper extremity disorders,[57] major consensus on testing methods has not emerged. Clearly, standardisation of functional measures in occupational epidemiology is essential to provide a consistent database, for better comparisons across industries.

In addition to a core set of standardised questions, future research may involve worker specific assessments[58] to address an individual's work related health concerns. By focusing on the important issues, investigators may apply effective interventions.[59]

CONCLUSION
Similar to laboratory tests, which can measure upper and lower limits of normal, investigators need questionnaires that can measure a fuller spectrum of function. Specifically, self reported measures are needed to assess mild to moderate disorders among workers to avoid more severe (and costly) consequences. Questionnaires in clinical research confer health benefits of detecting mild to moderate disorders. Early intervention with such disorders may allow a quicker return to normal function.

Functional status measures can quantify the impact of health on performance. Quantifying function may be useful in detecting mild to moderate disorders of the neck and upper extremity. Unfortunately, use of functional outcomes in epidemiological studies has been limited by lack of standardisation, and insufficient availability of normative data. Additionally, there is no gold standard for testing many upper extremity conditions.

Three measures were identified as most relevant for epidemiological studies among workers with mild to moderate upper extremity conditions: (1) the Nordic Musculoskeletal Questionnaire; (2) the Upper Extremity Questionnaire; and (3) the Neck and Upper Limb Instrument. Other functional measures may be relevant; however, they have not been tested successfully in field studies among workers.

Use of standardised functional measures with a wider spectrum of health is regarded as a realistic enhancement for research of neck and upper extremity disorders in the workplace. An important aspect of development is successful testing of validity, repeatability, and responsiveness among active workers. Appropriate use of functional status questionnaires is imperative for a meaningful portrayal of health.

ACKNOWLEDGEMENTS

. . . .

Authors' affiliations
D F Salerno, C Copley-Merriman, T N Taylor, Pfizer, Inc., Ann Arbor Laboratories, 2800 Plymouth Road, Ann Arbor, MI 48105, USA
J Shinogle, R M Schulz, University of South Carolina, College of Pharmacy and School of Public Health, Coker Life Sciences Building, Columbia, SC 29208, USA

REFERENCES

1 **Stock SR**, Cole DC, Tugwell P, et al. Review of applicability of existing functional status measures to the study of workers with musculoskeletal disorders of the neck and upper limb. Am J Ind Med 1996;**29**:679–88.
2 **McHorney CA**. Health status assessment for adults: past accomplishments and future challenges. Ann Rev Public Health 1999;**20**:309–35.
3 **Salerno DF**. Validity and reliability of health measures for upper extremity conditions among keyboard operators. PhD Dissertation. University of Michigan, 1998.
4 **Palmer K**, Cooper C. Repeated movements and repeated trauma affecting the musculoskeletal system. In: Adams PH, Aw T-C, Cockcroft A, et al, eds. Hunter's diseases of occupations, 9th edn. London: Arnold, 2000:453–75.
5 **United States Department of Labor**. Workplace injuries and illnesses in 1999. Bureau of Labor Statistics, USDL 00-357, 2000.
6 **Punnett L**. Epidemiologic studies of physical ergonomic stressors and musculoskeletal disorders. In: Work-related musculoskeletal disorders: report, workshop summary, and workshop papers. Washington, DC: National Academy Press, 1999.
7 **Mocci F**, Serra A Corrias GA. Psychological factors and visual fatigue in working with video display terminals. Occup Environ Med 2001;**58**:267–71.
8 **Viikari-Juntura E**, Martikainen R, Luukkonen R, et al. Longitudinal study on work related and individual risk factors affecting radiating neck pain. Occup Environ Med 2001;**58**:345–52.
9 **Nahit ES**, Macfarlane GJ, Pritchard CM, et al. Short term influence of mechanical factors on regional musculoskeletal pain: a study of new workers from 12 occupational groups. Occup Environ Med 2001;**58**:374–81.
10 **Simeonsson RJ**, Lollar D, Hollowell J, et al. Revision of the international classification of impairments, disabilities, and handicaps: development issues. J Clin Epidemiol 2000;**53**:113–24.
11 **Guyatt GH**, Feeny DH, Patrick DL. Measuring health-related quality of life. Ann Int Med 1993;**118**:622–9.
12 **Tarlov AR**, Greenfield S, Nelson EC, et al. The Medical Outcomes Study. JAMA 1989;**262**:925–30.
13 **Kozma CM**, Reeder CE, Schulz RM. Economic, clinical, and humanistic outcomes: a planning model for pharmacoeconomic research. Clin Ther 1993;**15**:1121–32.
14 **Streiner DL**, Norman GR. Health measurement scales: a practical guide to their development and use. Oxford: Oxford University Press, 1991.
15 **DeVellis RF**. Scale development: theory and applications. London: Sage Publications, 1991.
16 **Cohen J**. A coefficient of agreement for nominal scales. Educ Psychol Meas 1960;**20**:37–46.
17 **Fleiss JL**, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. Educ Psychol Meas 1973;**33**:613–19.
18 **Fleiss JL**. Statistical methods for rates and proportions. New York: John Wiley and Sons, 2nd edn. 1981.
19 **Kramer MS**, Feinstein AR. Clinical biostatistics LIV. The biostatistics of concordance. Clin Pharmacol Ther 1981;**29**:111–23.
20 **Müller R**, Büttner P. A critical discussion of intraclass correlation coefficients. Stat Med 1994;**13**:2465–76.
21 **Cronbach LJ**. Coefficient alpha and the internal structure of tests. Psychometrika 1951;**16**:297–334.
22 **Nunnally JC**, Bernstein IH. Psychometric theory. New York: McGraw Hill, 3rd edn. 1994.
23 **Cohen J**. Statistical power analysis for the behavioral sciences. Hillsdale, NJ: Erlbaum, 1988.
24 **Liang MH**. Evaluating measurement responsiveness. J Rheumatol 1995;**22**:1191–2.
25 **Roach KE**, Budiman-Mak E, Songsiridej N, et al. Development of a Shoulder Pain and Disability Index. Arthritis Care Res 1991;**4**:143–9.
26 **L'insalata JC**, Warren RF, Cohen SB, et al. A self-administered questionnaire for the assessment of symptoms and function of the shoulder. J Bone Joint Surg Am 1997;**79-A**:738–48.
27 **Meenan RF**, Mason JH Anderson JJ, et al. AIMS2: the content and properties of a revised and expanded Arthritis Impact Measurement Scales health status questionnaire. Arthritis Rheum 1992;**35**:1–10.
28 **Levine DW**, Simmons BP, Koris MJ, et al. A self-administered questionnaire for the assessment of severity of symptoms and functional status in carpal tunnel syndrome. J Bone Joint Surg 1993;**75A**:1585–92.
29 **Katz JN**, Gelberman RH, Wright EA, et al. Responsiveness of self-reported and objective measures of disease severity in carpal tunnel syndrome. Med Care 1994;**32**:1127–33.
30 **Hudak PL**, Amadio PC, Bombardier C, et al. Development of an upper extremity outcome measure: The DASH (Disabilities of the Arm, Shoulder and Hand). Am J Ind Med 1996;**29**:602–8 (published erratum in Am J Ind Med 1996;**30**:372).
31 **MacDermid JC**, Richards RS, Donner A, et al. Responsiveness of the SF-36, DASH, patient-related wrist evaluation and physical impairments in evaluating recovery after a distal radius fracture. J Hand Surg 2000;**25A**:330–40.
32 **Institute for Work & Health**. Development of the DASH, http://www.iwh.on.ca/Pages/Research/dash.htm, 2000.
33 **Chung KC**, Pillsbury MS, Walters MR, et al. Reliability and validity testing of the Michigan Hand Outcomes Questionnaire. J Hand Surg 1998;**23**:575–87.
34 **Chung KC**, Hamill JB, Walters MR, et al. The Michigan Hand Outcomes Questionnaire (MHQ): assessment of responsiveness to clinical change. Ann Plast Surg 1999;**42**:619–22.
35 **Swiontkowski MF**, Engelberg R, Martin DP, et al. Short musculoskeletal function assessment questionnaire: validity, reliability, and responsiveness. J Bone Joint Surg 1999;**81-A**:1245–60.
36 **Martin DP**, Engelberg R, Agel J, et al. Development of a musculoskeletal extremity health status instrument: The Musculoskeletal Function Assessment instrument. J Orthop Res 1996;**14**:173–81.
37 **Engelberg R**, Martin DP, Agel J, et al. Musculoskeletal Function Assessment instrument: criterion and construct validity. J Orthop Res 1996;**14**:182–92.
38 **Vernon H**, Mior S. The Neck Disability Index: a study of reliability and validity. J Manipulative Physiol Ther 1991;**14**:409–15.
39 **Hains F**, Waalen J, Mior S. Psychometric properties of the Neck Disability Index. J Manipulative Physiol Ther 1998;**21**:75–80.
40 **Salén BA**, Spangfort EV, Nygren ÅL, et al. Disability Rating Index: an instrument for the assessment of disability in clinical settings. J Clin Epidemiol 1994;**47**:1423–34.
41 **Pransky G**, Feuerstein M, Himmelstein J, et al. Measuring functional outcomes in work-related upper extremity disorders: development and validation of the Upper Extremity Function Scale. J Occup Environ Med 1997;**39**:1195–202.
42 **Kuorinka I**, Jonsson B, Kilbom Å, et al. Standardised Nordic questionnaires for the analysis of musculoskeletal symptoms. Appl Ergon 1987;**18**:233–7.
43 **Dickinson CE**, Campion K, Foster AF, et al. Questionnaire development: an examination of the Nordic Musculoskeletal Questionnaire. Applied Ergon 1992;**23**:197–201.
44 **Baron S**, Hales T, Hurrell J. Evaluation of symptom surveys for occupational musculoskeletal disorders. Am J Ind Med 1996;**29**:609–17.
45 **Franzblau A**, Werner R, Valle J, et al. Workplace surveillance for carpal tunnel syndrome: a comparison of methods. J Occup Rehab 1993;**3**:1–14.
46 **Franzblau A**, Salerno DF, Armstrong TJ, et al. Test-retest reliability of an upper extremity discomfort questionnaire in an industrial population. Scand J Work Environ Health 1997;**23**:299–307.
47 **Salerno DF**, Franzblau A, Armstrong TJ, et al. Test-retest reliability of the Upper Extremity Discomfort Questionnaire among keyboard operators. Am J Ind Med 2001;**40**:655–66.
48 **Stock SR**, Streiner D, Reardon R, et al. The impact of neck and upper limb disorders on the lives of affected workers: development of a new functional status index. Qual Life Res 1995;**4**:491.
49 **Physical Medicine Research Foundation**. A brief overview of the Neck and Upper Limb Index (NULI). http://www.icpro.org/ab_1_97.htm, 3 February 2000.
50 **Rose MS**, Koshman ML, Spreng S, et al. Statistical issues encountered in the comparison of health-related quality of life in diseased patients to

published general population norms: problems and solutions. *J Clin Epidemiol* 1999;**52**:405–12.

51 **Taylor RS**, Bonfiglio RP. Industrial rehabilitation medicine. Assessment of the outcome of treatment in industrial medicine, program development, documentation, and testimony. *Arch Phys Med Rehabil* 1992;**73**:S369–73.

52 **Kivimäki M**, Vahtera J, Ferrie JE, *et al*. Organisational downsizing and musculoskeletal problems in employees: a prospective study. *Occup Environ Med* 2001;**58**:811–17.

53 **Kreider B**. Latent work disability and reporting bias. *J Hum Resour* 1999;**34**:734–69.

54 **Ransohoff DF**, Feinstein AR. Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. *N Engl J Med* 1978;**299**:926–30.

55 **Salerno DF**, Franzblau A, Werner RA, *et al*. Median and ulnar nerve conduction studies among workers: normative values. *Muscle Nerve* 1998;**21**:999–1005.

56 **Armstrong TJ**, Buckle P, Fine LJ, *et al*. A conceptual model for work-related neck and upper-limb disorders. *Scand J Work Environ Health* 1993;**19**:73–84.

57 **Katz JN**, Stock SR, Evanoff BA, *et al*. Classification criteria and severity assessment in work-associated upper extremity disorders: methods matter. *Am J Ind Med* 2000;**38**:369–72.

58 **Wright JG**. Evaluating the outcome of treatment: shouldn't we be asking patients if they are better? *J Clin Epidemiol* 2000;**53**:549–53.

59 **Feinstein AR**, Josephy BR, Wells CK. Scientific and clinical problems in indexes of functional disability. *Ann Int Med* 1986;**105**:413–20.

ECHO ....

Downsizing in manufacturing can bring health benefits

Please visit the Occupational and Environmental Medicine website [www. occenvmed.com] for link to this full article.

A long term study of sawmilling in Canada has disclosed that workers who were re-employed outside the industry after its downsizing in the 1980s enjoyed better heath than those remaining. The odds of poor health was 1.47 (95% confidence interval 1.02 to 2.11) for those continuing to work in the mills compared with workers who were laid off and found work elsewhere.

During 1979–1998, 60% of mill workers lost their jobs. In 1998 about half of the workers under retirement age were employed in sawmills and half outside the industry.

Restructured mills employed far fewer unskilled workers, managers, and tradesmen but threefold more semiskilled workers. High strain jobs—more psychologically demanding, with no control—were reported to have decreased slightly, as were noise and physical demand.

More workers who were laid off experienced unemployment, more frequently, and for longer, than those who remained. By 1998 two thirds were working in the service sector and forest products. A higher proportion were managers (34%) than in the mills (3%) and significantly fewer were unskilled workers (18% *v* 43%). Control and social support scores were higher than in the mills, and physical demand and noise scores were lower.

The study was based on a sample of 3000 employees in 14 study sawmills who were drawn randomly from a 28 000 cohort. Data on employment, work conditions, and health were based on 1885 interviews in 1998. They were adjusted for current smoking, age, income, highest educational level, occupational category in 1979 and 1998, job strain, and periods of unemployment.

▲ *Journal of Epidemiology and Community Health* 2002;**56**:506–509.