AMERICAN JOURNAL OF INDUSTRIAL MEDICINE 38:369–372 (2000)

Commentary

Classification Criteria and Severity Assessment in Work-Associated Upper Extremity Disorders: Methods Matter

Jeffrey N. Katz, MD, MS,[1]* Susan R. Stock, MD, MSc, FRCPC,[2] Bradley A. Evanoff, MD, MPH,[3] David Rempel, MD,[4] J. Steven Moore, MD, MPH,[5] Alfred Franzblau, MD,[6] and Ronald H. Gray, MD, MSc[7]

Work-associated musculoskeletal disorders of the upper extremity are common and disabling. Research on these disorders is needed and requires valid methods of classification of the disorders for epidemiologic studies and measurement of their impact on functional status. This commentary discusses the methodologic aspects of classification and functional status assessment in upper extremity musculoskeletal disorders. Am. J. Ind. Med. 38:369–372, 2000. *© 2000 Wiley-Liss, Inc.*

KEY WORDS: upper extremity disorders; carpal tunnel syndrome; tendinitis; diagnosis; classification criteria; functional status

INTRODUCTION

Musculoskeletal disorders of the upper extremity were the fastest growing source of disability in the American workplace through the mid-1990s [Bernard, 1997]. Because the incidence, prevalence, and cost of these disorders are substantial, research is needed urgently. Specifically, we need refined estimates of prevalence, incidence, risk factors, prognostic factors, and impact, particularly in special populations such as older workers, minorities and students. Finally, we must define optimal strategies for prevention and medical and disability management, and develop improved health service delivery systems. Two fundamental methodologic issues are critical to this ambitious research agenda: development of classification criteria for epidemiological case definitions; and measurement of the severity of work-associated upper extremity disorders, particularly their impact on functional status.

These methodologic goals are challenging. The Center for Information Technology and Health Research at Johns Hopkins University convened a conference in November 1996 to address these issues. The conference activities included development of preliminary classification criteria for carpal tunnel syndrome [Rempel et al., 1998] and disorders of the distal upper extremity, as well as a critical review of the literature on functional status instruments

[1]Robert Brigham Multipurpose Arthritis and Musculoskeletal Diseases Center, Division of Rheumatology, Immunology and Allergy, Brigham and Women's Hospital, Harvard Medical School
[2]Montreal Department of Public Health and Department of Epidemiology, Biostatistics and Occupational Health, McGill University (SRS)
[3]Division of General Medical Sciences, Department of Medicine, Washington University School of Medicine (BAE)
[4]Department of Medicine, University of California, San Francisco (DR)
[5]Zachry Engineering Center, Texas A and M University (JSM)
[6]Department of Environmental and Industrial Health, University of Michigan School of Public Health (AF)
[7]Department of Population and Family Health Sciences, Johns Hopkins School of Hygiene and Public Health (RHG)
Contract grant sponsor: Center for Information Technology and Health Research, Johns Hopkins University.
*Correspondence to: Jeffrey N. Katz, MD, MS, Division of Rheumatology, Immunology and Allergy, Brigham and Women's Hospital, B-3, 75 Francis Street, Boston, MA 02115. E-mail: Jnkatz@bics.bwh.harvard.edu

Accepted 4 April 2000

relevant to work-associated neck and upper extremity disorders. Readers may contact the Center for a full report of the conference [Katz et al., 1998].

The development of criteria for classification or diagnosis would be straightforward if a gold standard diagnostic test existed. In medicine, a provisional diagnosis is frequently based upon a syndrome consisting of a cluster of symptoms and signs. Such syndromes are, however, rarely pathognomonic and require further investigation to provide a definitive diagnosis. Unfortunately, there is no criterion standard for any of the upper extremity soft tissue musculoskeletal conditions. This situation is not unique. In fact, there is no gold standard either for myocardial infarction, congestive heart failure, chronic obstructive lung disease, asthma, or a variety of other clinical syndromes. Clinicians have always had to define syndromes based upon an underlying clinical impression and then develop operational criteria using features of the history, physical examination, and laboratory evaluation. The challenge faced by researchers investigating musculoskeletal disorders of the upper extremity is particularly daunting, in part because the underlying relationships between symptoms, physical examination findings and test results (such as nerve conduction studies in carpal tunnel syndrome) are weak [Levine et al., 1993; Homan et al., 1999]. While a gold standard does not exist for myocardial infarction, the electrocardiogram, creatine phosophokinase, and troponin tests are more effective at distinguishing cases of myocardial infarction than any single objective measure in musculoskeletal disorders.

Case definitions for upper extremity musculoskeletal disorders are also problematic because of conceptual inconsistencies in the traditional nomenclature for these disorders. For example ''cumulative trauma disorder'' implies an understanding of etiology and pathogenesis; ''rotator cuff tendinitis'' implies an understanding of the anatomic defect; and ''impingement syndrome'' implies an understanding of pathophysiology. Because legitimate debate persists over the etiology, pathogenesis, anatomy and pathophysiology of these disorders, these terms may be less preferable than simple phenotypic descriptions such as ''shoulder pain,'' pedestrian as they may be. In back pain, the Quebec Task force on Spinal Disorders, in a large consensus process and systematic critical review of the literature, described similar limits of the diagnostic terms and nomenclature used for soft tissue spinal disorders. The Task Force members proposed a simplified classification system based primarily on symptoms, signs and, in the case of a few classes, imaging techniques [Atlas et al., 1996; Spitzer et al., 1987]. No such exercise has as yet been carried out for soft tissue upper extremity disorders. The 1996 Conference of the Center for Information Technology and Health Research was an initial attempt to address these issues.

Studies of classification criteria suggest that expert clinicians can more accurately identify cases than most history, physical examination, or laboratory parameters [Katz et al., 1990]. This observation supports the concept that, for some disorders, the best proxy for a gold standard may be the opinion of an expert clinician. Under this assumption, the development of classification criteria becomes an exercise in determining which history, physical examination, and laboratory findings match the impression of an experienced clinician. This exercise can be done empirically, as described below, or by using a Delphi or other consensus method. The consensus approach is often an efficient initial form of validation. Examples of the consensus approaches to case definition include the guidelines for occupational asthma developed by the American College of Chest Physicians [Chan-Yeung, 1995] and by the Canadian Thoracic Society [Tarlo et al., 1998].

The empirical approach has a long tradition in the field of rheumatology that concerns itself with complex clinical syndromes that do not have gold standard case definitions such as rheumatoid arthritis and systemic lupus erythematosus. In developing classification criteria for these disorders [Tan et al., 1982; Arnett et al., 1988], expert rheumatologists have first identified candidate history, physical examination, and laboratory findings that identify cases. Then they have assembled panels of patients including some they believe to be cases and others to be non-cases. They have administered the candidate history, physical examination, and laboratory evaluations and identified the findings that best distinguish cases. Recursive partitioning or regression procedures are useful techniques for identifying the best criteria sets or strategies for case definition.

Though seemingly simple, this exercise has important methodologic complexities. For the panel's results to be acceptable, the panel must be comprised of all potential stakeholders. At a minimum this would include a range of clinicians from different disciplines. For work-associated musculoskeletal disorders, the stakeholder group should ultimately be broadened to include the range of people who rely upon these syndromic diagnoses including Workers' Compensation Board representatives, industry, labor, etc.

Empirical validation must be carried out on suitable populations. The sensitivity and specificity of classification criteria vary according to the spectrum of disease in the sample studied [Ransohoff and Feinstein, 1978]. Thus, if criteria are to be used on workers, the empirical validation should be carried out in working populations, and should not be limited to hospital-based clinical populations. In fact, for criteria to be most robust several different validation exercises should be carried out in a range of patient samples. Thus, the process of validating and refining classification criteria spans years and includes many studies. Furthermore, as technology and knowledge of these disorders improve, the criteria must be updated to

reflect these advances or clinicians will regard them as outmoded.

The first critical step in the process of developing classification criteria is the nomination of salient history, physical examination, and laboratory parameters. The 1996 Conference at Johns Hopkins' accomplished this for carpal tunnel syndrome and disorders of the distal upper extremity, and proposed preliminary criteria sets. The empirical validation work has yet to be done; we hope some readers will take on such studies.

These exercises in developing classification criteria may seem unnecessarily complex, but they have substantial policy implications. Accurate classification criteria should lead ultimately to more precise estimates of the prevalence and burden of these conditions in working populations, which in turn may drive regulation, legislation, and funding priorities. It is difficult to predict the effect of more accurate classification criteria, such as those recently proposed for carpal tunnel syndrome [Rempel et al., 1998], on the management of individual workers such as decisions regarding treatments or disability awards. This point raises the distinction between classification and diagnosis. Classification criteria are generally used for research, and necessarily are practical, affordable, and brief. In contrast, management of individual workers with upper extremity disorders may at times require sophisticated diagnostic methods. Nerve conduction testing may be impractical for many population-based studies, but useful for identifying carpal tunnel syndrome in an individual worker.

Health status measures present a different set of methodologic challenges [Guyatt et al., 1993; Jette and Jette, 1996; Wood-Dauphinee, 1999]. Substantial research over the last two decades has demonstrated that elements of health-related quality of life (such as symptoms, physical function, social and role function, mental health) are the outcomes of greatest interest to patients with musculoskeletal disorders [Stock et al., 1996]. Health status measures are often more reproducible [Feinstein et al., 1986] and can better discriminate levels of severity, predict subsequent disability, and reflect clinical improvement or deterioration than traditional objective parameters such as physical examination or laboratory measures [Pincus et al., 1989; Katz et al., 1994]. The health status measures are also valid in working populations and recipients of Workers Compensation involved in research studies [Katz et al., 1996]. Aside from these psychometric advantages, functional status measures have the additional logistical appeal that they can be completed rapidly by patients either in person, telephone interviews, or mailed questionnaires. Finally, as a fundamental policy matter, functional improvement is a principal goal of care for injured workers and should be measured directly.

In the non-occupational literature it has been shown that disease-specific measures of functional status are generally more discriminating and responsive to change than generic measures [Bombardier et al., 1995; Amadio et al., 1996; Bessette et al., 1998]. Generic measures, however, facilitate comparisons across studies of different conditions. Many investigators advocate using both a disease-specific measure and a generic measure. Several recently developed measures of health-related quality of life are appropriate for work-associated neck and upper extremity disorders. Some are specific to particular diseases such as carpal tunnel syndrome [Levine et al., 1993] or shoulder problems [Roach et al., 1991] while others have been developed to encompass disorders throughout the upper extremity [Stock et al., 1995; Hudak et al., 1996]. Two measures [Stock et al., 1995; Pransky et al., 1997] were designed specifically for use in workers. In general these measures have been shown to be reliable, valid, and sensitive to change.

Researchers studying work-associated upper extremity syndromes are advised to choose among existing measures rather than develop new ones, as several existing tools appear adequate for most study settings, and development of new measures requires substantial investment. The choice of instrument should be based on the nature of the research or clinical question, the range and appropriateness of the domains of the instrument, the methods and population upon which development was based, and the demonstrated reproducibility, responsiveness, and validity relevant to the proposed application. If researchers feel that a new measure is needed, they should adhere to established methods for the development of quality of life measures. This requires careful attention to specification of appropriate domains, item selection, item reduction, formatting and pretesting of the questionnaire, and testing in appropriate samples for reproducibility, responsiveness, and validity [Guyatt et al., 1986, 1993; Jaeschke and Guyatt, 1990; Streiner and Norman, 1995; Testa and Simonson, 1996].

While health status assessment is a relatively mature area of investigation, there remains a formidable research agenda. First, workers tend to have little time to spare on lengthy surveys. This highlights the tradeoffs between psychometric advantages of lengthier questionnaires and the logistical benefits of shorter ones. More work is needed on developing yet briefer psychometrically sound instruments. Further work is also needed on more specific measures of functioning across a wide range of dimensions associated with the workplace [Lerner et al., 1997]. Evaluation of the distribution of scores in different workplace and rehabilitation contexts would help to elucidate the floor and ceiling effects of these instruments and assist in the interpretation of the scores. Valid measures of work-related utility and function would facilitate economic analyses that incorporate workers' quality of life as well as costs, including productivity. Some of these research areas are currently under investigation; we look forward to the fruits of these efforts.

REFERENCES

Amadio PC, Silverstein MD, Ilstrup DM, Schleck CD, Jensen LM. 1996. Outcome assessment for carpal tunnel surgery: the relative responsiveness of generic, arthritis-specific, disease-specific and physical examination measures. J Hand Surg 21A:338–346.

Arnett FC, Edworthy SM, Block DA, McShane DJ, Fries JF, Cooper NS, Healey LA, Kaplan SR, Liang MH, Luthra HS, Medsger TA Jr, Mitchell DM, Neustadt DH, Pinals RS, Schaller JG, Sharp JT, Wilder RL, Hunger GG. 1988. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 31:315–324.

Atlas SJ, Deyo RA, Patrick DL, Convery K, Keller RB, Singer DE. 1996. The Quebec Task Force classification for Spinal Disorders and the severity, treatment, and outcomes of sciatica and lumbar spinal stenosis. Spine 21:2885–2892.

Bernard BP, Editor. 1997. Musculoskeletal disorders and workplace factors. National Institute of Occupational Safety and Health, DHHS (NIOSH) Publication No. 97-141, Cincinnati, OH.

Bessette L, Sangha O, Kuntz KM, Keller RB, Lew RA, Fossel Ah, Katz JN. 1998. Comparative responsiveness of generic versus disease-specific and weighted versus unweighted health status measures in carpal tunnel syndrome. Med Care 36:491–502.

Bombardier C, Melfi C, Paul J, Green R, Hawker G, Wright J, Coyte P. 1995. Comparison of a generic and a disease-specific measure of pain and physical function after knee replacement surgery. Med Care 33:AS131–AS144.

Chan-Yeung M. 1995. American College of Chest Physicians Consensus Statement: Assessment of asthma in the workplace. Chest 108:1084–1117.

Feinstein AR, Josephy BR, Wells CK. 1986. Scientific and clinical problems in indexes of functional disability, Ann Intern Med 105: 413–420.

Guyatt GH, Bombardier C, Tugwell PX. 1986. Measuring disease-specific quality of life in clinical trials. Can Med Assoc J 134:889–895.

Guyatt GH, Feeny DH, Patrick DL. 1993. Measuring health-related quality of life. Ann Int Med 118:622–629.

Homan MM, Franzblau A, Werner RA, Albers JW, Armstrong TJ, Bromberg MB. 1999. Agreement between symptom surveys, physical examination procedures and electrodiagnostic findings for the carpal tunnel syndrome. Scan J Work Environ Health 25:115–124.

Jaesche R and Guyatt GH. 1990. How to develop and validate a new quality of life instrument. In Spilker B (ed.), Quality of life assessments in clinical trials, New York: Raven Press.

Jette DU and Jette AM. 1996. Health status assessment in the occupational health setting. Orthop Clin North Am 27:891–363.

Hudak PL, Amadio PC, Bombardier C, Upper Extremity Collaborative Group (UECG). 1996. Development of an upper extremity outcome measure: The DASH (Disabilities of the arm, shoulder, and hand). Ame J Indus Med 29:602–608.

Katz JN, Gelberman RH, Wright EA, Lew RA, Liang MH. 1994. Responsiveness of self-reported and objective measures of disease severity in carpal tunnel syndrome. Med Care 32:1127–1133.

Katz JN, Larson MG, Sabra A, Krarup C, Stirrat CR, Sethi R, Eaton HM, Fossel AH, Liang MH. 1990. The Carpal Tunnel Syndrome: diagnostic utility of the history and physical examination findings. Ann Intern Med 112:321–327.

Katz JN, Punnett L, Simmons BP, Fossel AH, Mooney N, Keller RB. 1996. Workers' compensation recipients with carpal tunnel syndrome: The validity of self-reported health measures. Am J Pub Health 86: 52–56.

Katz JN, Stock SR, Evanoff BA, Rempel D, Moore JS, Franzblau A, Gray RH. 1998. Working Paper: Conference on diagnostic criteria and severity assessment in work associated upper extremity disorders. Available through: Center for Information Technology and Health Research, Johns Hopkins School of Hygiene and Public Health, 615 North Wolfe Street, Baltimore, MD 21205.

Lerner DJ, Amick BC, Rogers Malspeis, Rogers WH, Gomes DRJ, Salem DN. 1997. The angina-related work limitations questionnaire. J Qual Life Res 7:23–32.

Levine DW, Simmons BP, Koris MJ, Daltroy LH, Hohl GG, Fossel AH, Katz JN. 1993. A self administered questionnaire for the assessment of severity of symptoms and functional status in carpal tunnel syndrome. J Bone Joint Surg 75A:1585–1592.

Pincus T, Callahan LF, Brooks RH, Fuchs HA, Olsen NJ, Kaye JJ. 1989. Self report questionnaire scores in rheumatoid arthritis compared with traditional physical, radiographic and laboratory measures. Ann Intern Med 110:259–266.

Pransky G, Feuerstein M, Himmelstein J, Katz JN, Vickers-Laht M. 1997. Measuring functional outcomes in work-related upper extremity disorders. Development and validation of the Upper Extremity Function Scale. J Occup Environ Med 39:1195–1202.

Ransahoff DF, Feinstein AR. 1978. Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. New Eng J Med 299: 926–930.

Rempel D, Evanoff B, Amadio PC, de Krom M, Franklin G, Franzblau A, Gray R, Gerr F, Hagbert M, Hales T, Katz JN, Pransky G. 1998. Consensus criteria for the classification of carpal tunnel syndrome in epidemiologic studies. Am J Public Health 88: 1447–1451.

Roach KE, Budiman-Mak E, Songsiridej N, Lertratanakul Y. 1991. Development of a shoulder pain and disability index. Arthr Care and Res 4:143–149.

Spitzer WO, LeBlanc FE, Dupuis M. 1987. Scientific approach to the assessment and management of activity-related spinal disorders. A monograph for clinicians. Report of the Quebec Task Force on Spinal Disorders. Spine 12:51–59.

Stock SR, Cole DC, Tugwell P, Streiner D. 1996. Review of applicability of existing functional status measures to the study of workers with musculoskeletal disorders of the neck and upper limb. Amer J Indus Med 29:679–688.

Stock SR, Streiner D, Reardon R, Darzins S, Dilworth P, Tugwell P, Loisel P. 1995. The impact of neck and upper limb disorders on the lives of affected workers: development of a new functional status index. Qua Life Res 4:491.

Streiner DL, Norman GR. 1995. Health measurement scales: A practical guide to their development and use, 2nd ed., New York: Oxford University Press.

Tan EM, Cohen AS, Fries JF, Masi AT, McShane DJ, Rothfield NF, Schaller JG, Talal N, Winchester RJ. 1982. The 1982 revised criteria for the classification of systemic lupus erythematosus (SLE). Arthr Rheum 25:1271–1277.

Tarlo SM, Boulet LP, Cartier A, Cockroft D, Cote J, Hargreave FE, Holness L, Liss G, Malo JL, Chan-Yeung M. 1998. Canadian Thoracic Society guidelines for occupational asthma. Can Respir J 5:397–410.

Testa MA, Simonson DC. 1996. Assessment of quality-of-life outcomes. New Eng J Med 334:835–840.

Wood-Dauphinee S. 1999. Assessing quality of life in clinical research: from where have we come and where are we going? J Clin Epidemiol 52:355–363.