<table>
<thead>
<tr>
<th>Page</th>
<th>Section</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>335</td>
<td>Commentary</td>
<td>Rehabilitation Medicine: The Asian Perspective.</td>
<td>Tai Ryoon Han, MD, PhD; Moon Suk Bang, MD, PhD</td>
</tr>
<tr>
<td>339</td>
<td>CME Article * 2007 Series * Number 1</td>
<td>Long-Term Survival in Werdnig-Hoffmann Disease.</td>
<td>John R. Bach, MD; Katherine Saltstein, MD; Diane Sinquee, MD; Brian Weaver, RT; Eugene Komaroff, PhD</td>
</tr>
<tr>
<td>346</td>
<td>Research Article</td>
<td>CME Self-Assessment Exam.</td>
<td></td>
</tr>
<tr>
<td>349</td>
<td>Research Article</td>
<td>Medical Considerations of Long-Term Survival of Werdnig-Hoffmann Disease.</td>
<td>John R. Bach, MD</td>
</tr>
<tr>
<td>356</td>
<td>Research Article</td>
<td>The Impact of the Inpatient Rehabilitation Facility Prospective Payment System on Stroke Program Outcomes.</td>
<td>Robert Gillen, PhD; Howard Tennen, PhD; Tara McKee, PhD</td>
</tr>
<tr>
<td>364</td>
<td>Research Article</td>
<td>The Clinician Effect on &quot;Objective&quot; Technical Components of the Electrodiagnostic Consultation.</td>
<td>K S.J. Yamakawa, MS; Andrew J. Haig, MD; M E. Geisser, PhD; H C. Tong, MD, MS; A Chiodo, MD; J A. Miner, MBA</td>
</tr>
<tr>
<td>373</td>
<td>Research Article</td>
<td>Fatigue-Induced Changes in Phasic Muscle Activation Patterns During Dynamic Trunk Extension Exercise.</td>
<td>Brian C. Clark, PhD; Todd M. Manini, PhD; Lori L. Ploutz-Snyder, PhD</td>
</tr>
<tr>
<td>379</td>
<td>CME Self-Assessment Exam</td>
<td>CME Self-Assessment Exam.</td>
<td></td>
</tr>
</tbody>
</table>
Research Article
Spine

380 A Clinical Tool for Office Assessment of Lumbar Spine Stabilization Endurance: Prone and Supine Bridge Maneuvers.
Kerri L. Schellenberg, MD; J Michael Lang, MD, FRCPC; K Ming Chan, MD, FRCPC;
Robert S. Burnham, MD, MSc, FRCPC

Research Article
Spinal Cord Injury

387 Factors Associated with Health-Related Quality of Life in Chronic Spinal Cord Injury.
Nitin B. Jain, MD, MSPH; Marianne Sullivan, PhD; Lewis E. Kazis, ScD; Carlos G. Tun, MD;
Eric Garshick, MD, MOH

Research Article
Referred Pain

397 Dry Needling to a Key Myofascial Trigger Point May Reduce the Irritability of Satellite MTrPs.
Yueh-Ling Hsieh, PhD, PT; Mu-Jung Kao, MD; Ta-Shen Kuan, MD, MS; Shu-Min Chen, MD;
Jo-Tong Chen, MD, PhD; Chang-Zern Hong, MD

Research Article
Spasticity

Thiru Annaswamy, MD, MA; Srinivas Mallempati, MD; Stephen C. Allison, PT, PhD;
Lawrence D. Abraham, EdD

Review & Analysis
Spasticity

412 Motor Neuron Inhibition-Based Gene Therapy for Spasticity.
Shearwood McClelland III, MD; Qingshan Teng, MD; Lyndsey S. Benson, BS; Nicholas M.
Boulis, MD

Visual Vignette

422 Spinal Injury Mimicking Lumbosacral Plexopathy.
Jennie-Corinne Baublitz-Brenenborg, DO; Michael J. DePalma, MD; David X. Cifu, MD
ABSTRACT


Key Words: Rehabilitation Medicine, Asia, Collaboration, Medical Societies

In this article, the authors discuss the history, current status, and future facing physical and rehabilitation medicine (PRM) in Asia, which, in this article, does not include the Middle East. The International Society of Physical and Rehabilitation Medicine (ISPRM), the most widely recognized rehabilitation medicine (RM) society, divides the world into seven regions and includes the Middle East with Africa.1

In contrast to developed countries, health care and welfare service resources are more limited in most Asian countries. Malnutrition, infectious disease, traffic and labor accidents, natural disasters, and even wars continue to play important causative roles of disabilities in many Asian countries. As a result, community-based rehabilitation has been adopted and is now practiced as a rehabilitation service strategy by many countries.2 Moreover, recent economic progress has elevated some nations to developed country status, and these countries now have advanced health care systems. Thus, the present status of RM in the Asian region represents a broad spectrum of medical rehabilitation services, ranging from the standards of developed countries to those of developing countries. RM in the Asian region is certain to progress in diverse ways that reflect the dynamics of Asian economies.

DATA COLLECTION

To obtain data on the current status of RM, a structured questionnaire was sent to the RM representatives of 15 Asian countries (China, Hong Kong, India, Indonesia, Japan, Korea, Laos, Malaysia, Mongolia, Nepal, the Philippines, Singapore, Taiwan, Thailand, and Vietnam). The questionnaire contained questions on general demographics, number of physicians and specialists of RM, history of RM, education, specialty training, CME, and main areas of clinical practice. All the representatives of 15 countries responded to the questionnaire and their names are listed in the Acknowledgments section below. For the data of Bangladesh, authors consulted the homepage of ISPRM.1
DEFINITION

The official title of the specialty differs in countries and is referred to as RM, PRM, or physical medicine and rehabilitation (PM&R). However, the term RM is more commonly used; this may reflect the greater priority given to medical rehabilitation over physical medicine in Asia, although the use of PRM is gaining ground because of the influence of the ISPRM, which was formed as the result of a merger between the International Rehabilitation Medicine Association and the International Federation of Physical and Rehabilitation Medicine in 1999.1

RM: Japan, Korea, The Philippines, Hong Kong, Malaysia, Singapore, Laos, Thailand
RM or PM&R: Vietnam, China
PM&R: Bangladeshi, India, Taiwan
PRM: Indonesia, Mongolia
Not decided: Nepal

HISTORY

National Organizations

In most Western countries, RM was started after World War II to provide medical service and rehabilitation for veterans who had experienced spinal cord injuries or amputations.3

RM was started much later in Asia than in Europe or the United States. In the Philippines and Japan, national organizations were started in the 1960s, and in the 1970s, Taiwanese, Korean, and Indian national organizations were established, followed by China, Indonesia, and Thailand in 1980s and by Vietnam, Hong Kong, and Laos in the 1990s. More recently, national organizations were established in Malaysia, Singapore, and Mongolia. In China and Thailand, two national organizations exist. In China, the Chinese Association of RM and the Chinese Society of Physical Medicine and Rehabilitation (CSPMR) coexist, though the specialty program in China falls under the auspices of the CSPMR (Table 1).

Regional Organizations

The first regional meeting in Asia was the first Asian Rehabilitation Medicine Association Congress, which was held in 1998 in Chiangmai, Thailand. The second meeting was hosted by the Philippines in 2002, with only few attendees from Asian countries, and it was the last meeting of the Asian Rehabilitation Medicine Association.

In 2001, the new millennium Asian Symposium on Rehabilitation Medicine was held in Tokyo, sponsored by the Japanese Association of RM. Representatives of 14 countries shared data and information on RM. As a result, a special report was published by Chino et al.2

In 2002, the biennial joint conference, organized by the Japanese Association of RM and the Korean Academy of RM was held. The first meeting was held in Gyeongju, Korea and consisted of

<table>
<thead>
<tr>
<th>Country</th>
<th>National Organization</th>
<th>Established in</th>
<th>Regular Meeting</th>
<th>Member of ISPRM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bangladesh</td>
<td>Bangladesh Association of PRM (BAPMR)</td>
<td>1995</td>
<td>n/a</td>
<td>Yes</td>
</tr>
<tr>
<td>China</td>
<td>Chinese Association of RM (CARM)</td>
<td>1983</td>
<td>Annual</td>
<td>Yes</td>
</tr>
<tr>
<td>China</td>
<td>Chinese Society of PM&amp;R (CSPMR)</td>
<td>1985</td>
<td>Annual</td>
<td>Yes</td>
</tr>
<tr>
<td>Hong Kong</td>
<td>Hong Kong Association of RM (HKARM)</td>
<td>1996</td>
<td>Annual</td>
<td>Yes</td>
</tr>
<tr>
<td>India</td>
<td>Indian Association of PM&amp;R (IAPMR)</td>
<td>1972</td>
<td>Very often</td>
<td>No</td>
</tr>
<tr>
<td>Indonesia</td>
<td>Indonesian of PM&amp;R Association (IPMR)</td>
<td>1987</td>
<td>Annual</td>
<td>No</td>
</tr>
<tr>
<td>Japan</td>
<td>Japanese Association of RM (JARM)</td>
<td>1963</td>
<td>Annual</td>
<td>Yes</td>
</tr>
<tr>
<td>Korea</td>
<td>Korean Academy of RM (KARM)</td>
<td>1972</td>
<td>Biannual</td>
<td>Yes</td>
</tr>
<tr>
<td>Laos</td>
<td>None</td>
<td>n/a</td>
<td>n/a</td>
<td>No</td>
</tr>
<tr>
<td>Malaysia</td>
<td>Malaysian Association of RM (MARM)</td>
<td>2004</td>
<td>4 times/year</td>
<td>No</td>
</tr>
<tr>
<td>Mongolia</td>
<td>Mongolian Society of PRM (MSPRM)</td>
<td>2005</td>
<td>Biannual</td>
<td>Yes</td>
</tr>
<tr>
<td>Nepal</td>
<td>None</td>
<td>n/a</td>
<td>n/a</td>
<td>No</td>
</tr>
<tr>
<td>The Philippines</td>
<td>Philippine Academy of RM (PARM)</td>
<td>1960</td>
<td>Annual</td>
<td>Yes</td>
</tr>
<tr>
<td>Singapore</td>
<td>Society of RM Singapore (SRMS)</td>
<td>2005</td>
<td>5–6 times/year</td>
<td>Awaiting entry</td>
</tr>
<tr>
<td>Taiwan</td>
<td>Taiwan Academy of PM&amp;R (TAPMR)</td>
<td>1971</td>
<td>Annual</td>
<td>Yes</td>
</tr>
<tr>
<td>Thailand</td>
<td>Thai RM Association (TRMA)</td>
<td>1988</td>
<td>Annual</td>
<td>No</td>
</tr>
<tr>
<td>Thailand</td>
<td>Royal College of Physiatrists of Thailand (RCPT)</td>
<td>1998</td>
<td>Annual</td>
<td>Yes</td>
</tr>
<tr>
<td>Vietnam</td>
<td>Vietnam Rehabilitation Association (VINAREHA)</td>
<td>1991</td>
<td>Every 5 years</td>
<td>No</td>
</tr>
</tbody>
</table>

NA, not available; PMR, physical and rehabilitation medicine; PM&R, physical medicine and rehabilitation; RM, rehabilitation medicine.

plenary sessions, lectures, and free paper and poster sessions; the second meeting was held in Kyoto, Japan. The third meeting was postponed because of the scheduled 2007 ISPRM meeting in Seoul. However, it is hoped that the number of participating countries will increase at the next joint conference and that it will become established as a regional conference in the East Asian region.

In 2006, a meeting to establish a new regional organization for Asia and Oceania was held in Zhuhai, China. Representatives from 11 countries have agreed to form this organization, called the Asian Oceanian Society of PRM at the 2007 Seoul ISPRM meeting and to hold the first meeting in Nanzing, China in May 2008. Participants will include representatives from Australia, China, Hong Kong, India, Indonesia, Japan, Korea, Malaysia, Singapore, Thailand, and Vietnam.

SPECIALTY TRAINING AND BOARD CERTIFICATION

Ten countries (China, Hong Kong, India, Indonesia, Japan, Korea, Mongolia, the Philippines, Taiwan, and Thailand) have a specialty board examination and certification procedure. Singapore has a system similar to that of a board certification examination, with an exit examination after 3 yrs of training, including an interview and a structured clinical examination governed by the Academy of Medicine, Singapore. Malaysia has a similar university-based examination, with a board of examiners (external or international professional members). Bangladesh and Vietnam have specialty training programs but no national specialty board certification system; Nepal and Laos have a few doctors practicing as physiatrists, but no specialty training is available as yet.

Postgraduate specialty programs of 3–5 yrs are running in most countries. In Hong Kong, 6 yrs of training are required, whereas in Mongolia less than 1 yr is required. However, although the contents and levels of these specialty training programs vary among countries, standardization is expected in the near future.

The annual number of trainees is as high as 80–90 in Korea, followed by Taiwan, Japan, and the Philippines. Japan and Korea have CME systems with credits, and many countries have diverse educational systems like CME systems. The majority of countries regularly hold annual academic meetings, but the contents and frequency of these programs vary (Table 2). In terms of subspecialty training, questionnaire responses have shown that most countries would not favor an official

<table>
<thead>
<tr>
<th>Country</th>
<th>Population in Millions</th>
<th>No. of Physicians</th>
<th>No. of Physiatrists</th>
<th>No. of Trainees</th>
<th>Specialty Training yrs</th>
<th>Board Exam</th>
<th>Activities Like CME</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bangladesh</td>
<td>144.4</td>
<td>38,485</td>
<td>25</td>
<td>n/a</td>
<td>3</td>
<td>Yes</td>
<td>10 or more/year</td>
</tr>
<tr>
<td>China</td>
<td>1,283.60</td>
<td>1,520,967</td>
<td>Several thousands</td>
<td>n/a</td>
<td>2 or 3</td>
<td>Yes</td>
<td>Very often</td>
</tr>
<tr>
<td>India</td>
<td>1,119,50</td>
<td>71</td>
<td>39</td>
<td>n/a</td>
<td>3</td>
<td>Yes</td>
<td>Very often</td>
</tr>
<tr>
<td>Indonesia</td>
<td>1,250,50</td>
<td>416</td>
<td>25</td>
<td>n/a</td>
<td>4</td>
<td>Yes</td>
<td>Very often</td>
</tr>
<tr>
<td>Japan</td>
<td>1,251,689</td>
<td>11</td>
<td>25</td>
<td>n/a</td>
<td>4</td>
<td>Yes</td>
<td>Very often</td>
</tr>
<tr>
<td>Korea</td>
<td>1,251,689</td>
<td>11</td>
<td>25</td>
<td>n/a</td>
<td>4</td>
<td>Yes</td>
<td>Very often</td>
</tr>
<tr>
<td>Laos</td>
<td>1,251,689</td>
<td>11</td>
<td>25</td>
<td>n/a</td>
<td>4</td>
<td>Yes</td>
<td>Very often</td>
</tr>
<tr>
<td>Malaysia</td>
<td>1,251,689</td>
<td>11</td>
<td>25</td>
<td>n/a</td>
<td>4</td>
<td>Yes</td>
<td>Very often</td>
</tr>
<tr>
<td>Mongolia</td>
<td>1,251,689</td>
<td>11</td>
<td>25</td>
<td>n/a</td>
<td>4</td>
<td>Yes</td>
<td>Very often</td>
</tr>
<tr>
<td>Nepal</td>
<td>1,251,689</td>
<td>11</td>
<td>25</td>
<td>n/a</td>
<td>4</td>
<td>Yes</td>
<td>Very often</td>
</tr>
<tr>
<td>The Philippines</td>
<td>1,251,689</td>
<td>11</td>
<td>25</td>
<td>n/a</td>
<td>4</td>
<td>Yes</td>
<td>Very often</td>
</tr>
<tr>
<td>Singapore</td>
<td>1,251,689</td>
<td>11</td>
<td>25</td>
<td>n/a</td>
<td>4</td>
<td>Yes</td>
<td>Very often</td>
</tr>
<tr>
<td>Taiwan</td>
<td>1,251,689</td>
<td>11</td>
<td>25</td>
<td>n/a</td>
<td>4</td>
<td>Yes</td>
<td>Very often</td>
</tr>
<tr>
<td>Thailand</td>
<td>1,251,689</td>
<td>11</td>
<td>25</td>
<td>n/a</td>
<td>4</td>
<td>Yes</td>
<td>Very often</td>
</tr>
<tr>
<td>Vietnam</td>
<td>1,251,689</td>
<td>11</td>
<td>25</td>
<td>n/a</td>
<td>4</td>
<td>Yes</td>
<td>Very often</td>
</tr>
</tbody>
</table>

subspecialty program modeled on the U.S. subspecialty board systems for pediatric rehabilitation and spinal cord injury medicine. Some countries have established subspecialty training programs; for example, Taiwan has programs for soft-tissue echography, cardiac rehabilitation, pulmonary rehabilitation, geriatric rehabilitation, and cancer rehabilitation, and India has established fellowships and short-term training programs for gait and urorehabilitation.

**CLINICAL PRACTICE**

Current health care systems and RM practices depend on medical traditions and economic situations, but most countries show similar trends in the major fields. Brain disorders represent the largest area in inpatient and outpatient clinical practice in the majority of countries, followed by spinal cord injuries, musculoskeletal disorders, rheumatic disorders, pain, pediatric rehabilitation, cerebral palsy, and sports medicine. Research activity in RM is becoming active in China, Japan, Korea, and Taiwan in many specialty fields.

**FUTURE**

Some Asian countries do not have an official RM service system or offer only limited service. However, RM is growing rapidly in the region, and the need for RM will expand with industrialization, economic growth, and political maturity. The numbers of physiatrists and trainees are increasing rapidly in many countries compared with the data reported in 2001 (Table 2).

Regional meetings are expected to become more active, and it is expected that collaboration between countries via joint conferences, such as those currently held between Japan and Korea, will be held more frequently and will embrace clinical and research issues, thus promoting the quality of life for people with disabilities in Asia.

The 2007 ISPRM meeting will be held in Seoul, with the theme *West Meets East in Rehabilitation Medicine: New Challenges for a Better World*. This meeting is expected to be a turning point and to advance RM in many Asian developing countries and to introduce traditional Oriental medicine rehabilitation practices to countries in other regions. Thus, this meeting is expected to meaningfully contribute to the standardization and globalization of RM, and Asia will constitute the principal means of harmonizing global requirements.

**ACKNOWLEDGMENTS**

We thank Dr. Jianan Li, Dr. Tiebin Yan (China), Dr. Leonard S.W. Li (Hong Kong), Dr. Ajit Kumar Varma (India), Dr. Rosiana Pradasari Wirawan (Indonesia), Dr. Masami Akai (Japan), Dr. Bouathep Phoumindr (Laos), Dr. Zaliha Omar (Malaysia), Dr. Baljinnyam Avirmed (Mongolia), Dr. Ashok Bajracharya (Nepal), Dr. Reynaldo R. Rey-Matias (The Philippines), Dr. Karen Chua (Singapore), Dr. Chein-Wei Chang (Taiwan), Dr. Sukajan Pongprapai (Thailand), and Dr. Nguyen Xuan Nghien (Vietnam) for their help with the data collection.

**REFERENCES**

CME Objectives: Upon completion of this article, the reader should be able to: (1) explain causes of morbidity and mortality in children with type 1 spinal muscular atrophy (SMA 1), (2) identify the benefits of tracheostomy and noninvasive ventilation on survival in children with SMA 1, and (3) identify complications of respiratory management in children with SMA 1.

Level: Advanced.

Accreditation: The Association of Academic Physiatrists is accredited by the Accreditation Council for Continuing Medical Education to provide continuing medical education for physicians. The Association of Academic Physiatrists designates this continuing medical education activity for a maximum of 1.5 credits in Category 1 of the Physician’s Recognition Award of the American Medical Association. Each physician should claim only those credits that he or she actually spent in the education activity.

Disclosures: This work was performed at University Hospital, Newark, NJ. This work has not previously been presented. Disclosure statements have been obtained regarding the authors’ relationships with financial supporters of this activity. There are no apparent conflicts of interests related to the context of participation of the authors of this article.

ABSTRACT


Objectives: To report long-term survival of spinal muscular atrophy type 1 (SMA 1) and consequences on speech and ventilator dependence as a function of mode of ventilator use.

Design: A retrospective chart review of 106 consecutively referred SMA 1 patients, the 92 most severe of which were considered in three groups: untreated (group 1), tracheostomy managed (group 2), and noninvasively managed (group 3).

Results: The untreated patients died at 9.6 ± 4.0 mos of age. The mean age of the 22 patients referred with tracheostomy tubes (group 2) was 70.5 ± 43.3 mos (range 2–159 mos); five died at 66.2 ± 114.2 mos (range 8–270 mos) of age. Six had comprehendible speech at the time of tracheotomy and retained some ability to vocalize afterward. None of the 21 patients who had not developed the ability to speak did so after tracheotomy. Twenty-five of the 27 total lost all autonomous breathing ability immediately, and definitively, after tracheotomy. The 47 patients who used noninvasive mechanical ventilation (NIV) (group 3) were extubated to it during episodes of acute respiratory failure. Thirty-nine of these were 65.2 ± 45.8 mos (range 11–153 mos) of age, and eight died at 60.9 ± 26.1 mos (range 36–111 mos) of age. There was no significant difference in longevity with or without tracheostomy, but the NIV patients had significantly fewer ($P = 0.04$) hospitalizations per year after age 5; 39 of the 47 could communicate verbally, and only nine were continuously dependent on NIV.

Conclusions: NIV and tracheostomy can both prolong survival for SMA 1 patients, but the latter results in continuous ventilator dependence and speech does not develop.

Key Words: Mechanical Ventilation, Respiratory Paralysis, Quality of Life, Spinal Muscular Atrophy Type 1
The spinal muscular atrophies (SMAs) are inherited as autosomal recessive disorders of anterior horn cells with the genetic defect at chromosome 5q13. Gene deletions are detectable in 98% of patients. The incidence is about 1/5000. Severity is inversely proportional to the amount of survival motor neuron protein present in the anterior horn cells. It ranges from essentially total paralysis and a need for ventilatory support from birth to muscle weakness first presenting in adults.

The SMAs have been separated arbitrarily into four clinical severities, 1 through 4. Children with SMA type 1 (SMA 1), also commonly known as Werdnig–Hoffmann disease, never attain the ability to sit independently. This is the most common inherited fatal disease in infants and has been reported to be fatal by 2 yrs of age, with 50% mortality by 7 mos and 90% mortality by 12 mos of age. All untreated children with SMA 1 have paradoxical breathing that results in pectus excavatum and a narrow, funnel-shaped chest. Although SMA 1 is defined by not attaining the ability to sit, about 10% of patients do not develop respiratory failure or can still take some food by mouth after their second birthdays; thus, they are atypically mildly affected. Upper respiratory tract infections or, occasionally, aspiration attributable to dysphagia or gastroesophageal reflux develops into pneumonia and respiratory failure largely because of an ineffective cough. Most physicians discourage endotracheal intubation and tracheotomy, feeling that the prognosis for survival would not be greatly improved and that quality of life is too poor to justify invasive interventions.

When children with SMA are intubated for respiratory failure, they conventionally are ventilator weaned before attempts are made at extubation, and then they are extubated with oxygen supplementation and, possibly, continuous positive airway pressure (PAP) or low-span bilevel PAP. Only 6% of such conventional extubation attempts for patients intubated for respiratory failure are successful. The small minority who undergo tracheotomy often survive age 2, but they can die from complications related to the tube.

In a previous publication we have reported that although intercurrent chest colds can necessitate intensive care and intubation, tracheotomy can be avoided for most SMA 1 children by using an extubation protocol in which, in addition to conventional medical therapy, respiratory therapy, and nutritional support, mechanical insufflation–exsufflation (In-exsufflator, J. H. Emerson Company, Cambridge, MA) is used via the translaryngeal tube to clear airway secretions and is then used via an oronasal interface after extubation, and patients are extubated to high-span nasal bilevel PAP (inspiratory muscle aid) only after the SpO₂ has remained above 94% in ambient air for at least 12 hrs. Subsequently, care providers are trained and equipped to use oximetry as feedback to use high-span bilevel PAP and mechanically assisted coughing (MAC) (expiratory muscle aid) to reverse any decreases in SpO₂ below 95% in ambient air. MAC is the use of mechanical insufflation–exsufflation (CoughAssist, J. H. Emerson Company, Cambridge, MA) with an exsufflation-timed abdominal thrust. The protocol extubation success rate was 85% by comparison with the conventional 6%. The purpose of this work is to update the survival benefits of the use of respiratory muscle aids.

MATERIALS AND METHODS

The statuses of all 106 SMA 1 patients consecutively referred to a neuromuscular disease clinic from 1993 to April 2006 were reviewed. One child had an equivocal diagnostic workup, and two were lost to follow-up after an initial visit before 12 mos of age and were eliminated from further study. SMA 1 was diagnosed on the basis of DNA evidence of chromosome 5 exon 7 and 8 deletion in 95 of 103 children, affected siblings with genetically confirmed disease in five patients, and characteristic laboratory, muscle biopsy, and electromyography results in three children who did not undergo DNA testing.

This work was approved by our institutional review board. Data were gathered from clinical history, physical examination, and follow-up telephone calls in June 2006. Data from the children of the 11 families not contacted by telephone in June
interface was placed and the PAP was initiated during described sleep high-span bilevel PAP. The nasal intratracheostomy tubes at initial referral were prescribed continuous respiratory support before 18 mos of age. Group 3 patients used NIV and were successfully extubated after extubation, ventilator weaning was not attempted at the expense of hypercapnia, and extubation was done, irrespective of the extent of ventilator dependence, to high-span bilevel PAP with no ongoing postextubation supplemental oxygen. Patients who underwent tracheotomy while intubated were intubated a total of 236 times, with 194 of these intubations being for respiratory failure and 42 for surgery.

The 92 patients are considered in three groupings. For group 1, 18 patients died from respiratory failure because noninvasive mechanical ventilation (NIV) and acute invasive respiratory support were refused. For group 2, 27 patients underwent tracheotomy, including 15 who had tracheostomy tubes before referral to our clinic. Five of the 27 patients were initially managed noninvasively and were trached after 24 mos of age. For group 3, 47 patients used NIV and were successfully extubated to NIV during intercurrent episodes of acute respiratory failure. All of the parents were trained how to NIV during intercurrent episodes of acute respiratory failure before 1 yr of age. These patients were intubated a total of 236 times, with 194 of these intubations being for respiratory failure and 42 for surgery.

Inclusion criteria for typical SMA 1 are noted in Table 1. Of the 92 patients, 79 had initial episodes of respiratory failure that necessitated continuous respirator use (by tube or noninvasively), and 57 had been intubated at least once for respiratory failure before 1 yr of age. These patients were intubated a total of 236 times, with 194 of these intubations being for respiratory failure and 42 for surgery.

Eleven patients developed respiratory failure or were able to receive nourishment by mouth after their second birthdays and were, therefore, atypically mild, so their outcomes were considered separately. Thus, the comparison of three management approaches was determined for only the 92 children with typical, severe SMA 1.

Inclusion criteria for typical SMA 1 are noted in Table 1. Of the 92 patients, 79 had initial episodes of respiratory failure that necessitated continuous respirator use (by tube or noninvasively), and 57 had been intubated at least once for respiratory failure before 1 yr of age. These patients were intubated a total of 236 times, with 194 of these intubations being for respiratory failure and 42 for surgery.

<table>
<thead>
<tr>
<th>TABLE 1 Inclusion criteria for “typical” spinal muscular atrophy type 1 for this study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inability to roll or sit unsupervised at any time</td>
</tr>
<tr>
<td>Paradoxical breathing</td>
</tr>
<tr>
<td>Never attaining a (cry or volitional) vital capacity of 250 ml</td>
</tr>
<tr>
<td>One or more episodes of respiratory failure requiring continuous respiratory support before 18 mos of age</td>
</tr>
<tr>
<td>Loss of ability to receive any nutrition by mouth before 24 mos of age</td>
</tr>
</tbody>
</table>

2006 were taken from their previous clinic evaluations.

Eleven patients developed respiratory failure or were able to receive nourishment by mouth after their second birthdays and were, therefore, atypically mild, so their outcomes were considered separately. Thus, the comparison of three management approaches was determined for only the 92 children with typical, severe SMA 1.

Inclusion criteria for typical SMA 1 are noted in Table 1. Of the 92 patients, 79 had initial episodes of respiratory failure that necessitated continuous respirator use (by tube or noninvasively), and 57 had been intubated at least once for respiratory failure before 1 yr of age. These patients were intubated a total of 236 times, with 194 of these intubations being for respiratory failure and 42 for surgery.

The 92 patients are considered in three groupings. For group 1, 18 patients died from respiratory failure because noninvasive mechanical ventilation (NIV) and acute invasive respiratory support were refused. For group 2, 27 patients underwent tracheotomy, including 15 who had tracheostomy tubes before referral to our clinic. Five of the 27 patients were initially managed noninvasively and were trached after 24 mos of age. For group 3, 47 patients used NIV and were successfully extubated to NIV during intercurrent episodes of acute respiratory failure. All of the parents were trained how to NIV during intercurrent episodes of acute respiratory failure before 1 yr of age. These patients were intubated a total of 236 times, with 194 of these intubations being for respiratory failure and 42 for surgery.

Inclusion criteria for typical SMA 1 are noted in Table 1. Of the 92 patients, 79 had initial episodes of respiratory failure that necessitated continuous respirator use (by tube or noninvasively), and 57 had been intubated at least once for respiratory failure before 1 yr of age. These patients were intubated a total of 236 times, with 194 of these intubations being for respiratory failure and 42 for surgery.

The 92 patients are considered in three groupings. For group 1, 18 patients died from respiratory failure because noninvasive mechanical ventilation (NIV) and acute invasive respiratory support were refused. For group 2, 27 patients underwent tracheotomy, including 15 who had tracheostomy tubes before referral to our clinic. Five of the 27 patients were initially managed noninvasively and were trached after 24 mos of age. For group 3, 47 patients used NIV and were successfully extubated to NIV during intercurrent episodes of acute respiratory failure. All of the parents were trained how to use MAC and high-span (inspiratory minus expiratory) PAP greater than 10 cm H2O) bilevel PAP, or volumetric ventilators for those with tracheostomy tubes and for most NIV users over 6 yrs of age. Oxygen supplementation, continuous PAP, and low-span bilevel PAP were avoided.

All group 2 and 3 patients were prescribed oximeters and MAC as needed, either via oronasal interface or tracheostomy tube. Group 3 patients and the 12 group 2 patients who did not have tracheostomy tubes at initial referral were prescribed sleep high-span bilevel PAP. The nasal interface was placed and the PAP was initiated during sleep by the parents in the home. The bilevel PAP was used continuously during chest infections at spans of 13–17 cm H2O, with inspiratory PAPs up to 22 cm H2O. Because these infants could not trigger bilevel PAP machines, to facilitate patient–ventilator synchrony, backup rates slightly higher than the children's spontaneous rates were used. As the deep insufflations increased tidal volumes and caused chest expansion rather than retraction, the spontaneous breathing rates decreased and the bilevel rates were decreased accordingly. The goal was to optimize chest expansion during sleep, to rest inspiratory muscles, and to prevent or reverse pectus excavatum. In addition, paradoxical perspiration, flushing, and frequent arousals were decreased or eliminated by treatment. Most children 6 mos and older used the Respironics Pediatric Petite Nasal Interface. Younger children used interfaces adapted from infant “CPAP” circuits and small adult nasal prongs as previously described. Usually within 1 wk of initiation, the children would wait to be placed on bilevel PAP before going to sleep.

Oximetry was used continuously during chest infections or other periods of airway congestion as feedback to guide in the use of high-span bilevel PAP and, especially, MAC. The positive pressure (insufflation) also was used daily via oronasal interfaces for children over 8 mos of age for deep lung and chest wall mobilization and to accustom the child to it to facilitate eventual cooperation with it. It was used up to every 5 mins during chest infections for airway mucus expulsion either via oronasal interfaces or tracheostomy tubes if present. The pressures used were usually at 35–40 cm H2O to −35 to −40 cm H2O, but the goal was for rapid, full chest expansion and then full lung emptying. Besides the use of broad spectrum antibiotics, the patients' parents were told to seek intensive medical attention when the SpO2 baseline could not be kept above 94% despite aggressive MAC and continuous use of high-span bilevel PAP.

After intubation for respiratory failure, extubations were considered successful when reintubation was not required during the current hospitalization. Exubtations were considered *conventional* when supplemental oxygen or CPAP was used after extubation; they were considered *conventional plus* when postextubation bilevel PAP was used, or *protocol* when patients were only extubated once SpO2 was normal (>94%) in ambient air, MAC was used via the tube before and via oronasal interface after extubation, ventilator weaning was not attempted at the expense of hypercapnia, and extubation was done, irrespective of the extent of ventilator dependence, to high-span bilevel PAP with no ongoing postextubation supplemental oxygen. Patients who underwent tracheotomy while intubated were intubated a total of 236 times, with 194 of these intubations being for respiratory failure and 42 for surgery.

Inclusion criteria for typical SMA 1 are noted in Table 1. Of the 92 patients, 79 had initial episodes of respiratory failure that necessitated continuous respirator use (by tube or noninvasively), and 57 had been intubated at least once for respiratory failure before 1 yr of age. These patients were intubated a total of 236 times, with 194 of these intubations being for respiratory failure and 42 for surgery.

The 92 patients are considered in three groupings. For group 1, 18 patients died from respiratory failure because noninvasive mechanical ventilation (NIV) and acute invasive respiratory support were refused. For group 2, 27 patients underwent tracheotomy, including 15 who had tracheostomy tubes before referral to our clinic. Five of the 27 patients were initially managed noninvasively and were trached after 24 mos of age. For group 3, 47 patients used NIV and were successfully extubated to NIV during intercurrent episodes of acute respiratory failure. All of the parents were trained how to use MAC and high-span (inspiratory minus expiratory) PAP greater than 10 cm H2O) bilevel PAP, or volumetric ventilators for those with tracheostomy tubes and for most NIV users over 6 yrs of age. Oxygen supplementation, continuous PAP, and low-span bilevel PAP were avoided.

All group 2 and 3 patients were prescribed oximeters and MAC as needed, either via oronasal interface or tracheostomy tube. Group 3 patients and the 12 group 2 patients who did not have tracheostomy tubes at initial referral were prescribed sleep high-span bilevel PAP. The nasal interface was placed and the PAP was initiated during
intubated were not considered extubation failures. No nonprotocol extubation attempts took place at our institution. Parents were taught the three extubation approaches and were questioned concerning the approaches used for their children at other institutions. The children’s physicians were questioned in this regard after failed extubations and before the patients’ transfers to our unit for protocol extubations.

Avoided hospitalizations were defined as episodes of fever, airways productive of mucopurulent sputum that caused oxyhemoglobin desaturations below 95% but that were normalized with use of MAC, and ventilatory failure managed at home by continuous dependence on high-span bilevel PAP with little or no autonomous breathing ability. Post hoc pairwise tests with Tukey–Kramer adjustment for multiple comparisons were used to compare the groups. Hospitalization rate ratios also were compared for statistical significance.

**RESULTS**

Among the 11 relatively mild SMA 1 patients, one died at 43 mos of age. This patient had been using high-span bilevel PAP 12 hrs/day since 14 mos of age. His first hospitalization for respiratory failure was at age 16 mos, but he did not require a gastrostomy tube until 28 mos of age. He had been hospitalized six times for respiratory failure. He died when arresting from airway mucus plugging during a chest infection in an ambulance. His parents had been interdicted from accompanying him, MAC was discontinued, and supplemental oxygen was provided in the ambulance. The mean age of the other 10 patients is 8 yrs 3 mos (range 46–205 mos). Their mean age at first hospitalization for respiratory failure was 22.5 ± 16.7 mos (range 10–37 mos). Two did not have gastrostomy tubes at 46 and 183 mos of age. The other eight had gastrostomy tubes placed at 52.1 ± 34.3 mos of age (range 6–179 mos). Their mean maximum recorded vital capacities (VCs) were 410 ± 530 ml (range 210–1300 ml). All of the patients use high-span bilevel PAP for sleep and, in three cases, including the patient whose vital capacity peaked at 1300 ml and is now 880 ml, require it for more than 16 hrs/day.

The ages at referral, equipment setup, surgery, current age, and death for the three “typical” SMA 1 groups are noted in Table 2. The group 1 (untreated) patients died significantly younger (P < 0.001) than patients from groups 2 and 3, but the age at which they were first intubated was not significantly different than for group 2 (P = 0.99) or group 3 (P = 0.49). Hospitalizations per patient-year are noted in Table 3.

Of the 27 group 2 (tracheostomy) patients, seven, ten, and ten children underwent tracheotomy during their first, second, and subsequent hospitalizations at 1–37 mos of age. This group was hospitalized significantly less (P = 0.66) until age 3 but significantly more (P < 0.04) after age 5. Five of the 27 are deceased. One patient died at 270 mos of age from a bradycardia; this patient had received continuous respiratory support via tracheostomy tube from 2 mos of age. Two others died suddenly. They had chronic purulent bronchitis associated with pathogenic bacterial colonization of the airway; thus, they died from complications related to the tubes. One patient died suddenly from persistent lung collapse attributable to congenital bronchomalacia, and one died 2 wks after a

<table>
<thead>
<tr>
<th>Age, mos</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Referral</td>
<td>18</td>
<td>6.6 ± 4.1</td>
<td>27</td>
<td>25.7 ± 44.3</td>
</tr>
<tr>
<td>First hospitalization</td>
<td>18</td>
<td>4.1 ± 3.8</td>
<td>27</td>
<td>6.1 ± 2.9</td>
</tr>
<tr>
<td>First intubation</td>
<td>18</td>
<td>7.0 ± 13.6</td>
<td>27</td>
<td>7.1 ± 5.9</td>
</tr>
<tr>
<td>Gastrostomy</td>
<td>27</td>
<td>7.7 ± 5.6</td>
<td>45</td>
<td>11.7 ± 9.1</td>
</tr>
<tr>
<td>Tracheostomy</td>
<td>27</td>
<td>14.8 ± 15.2</td>
<td>5</td>
<td>7.8 ± 4.2</td>
</tr>
<tr>
<td>BiPAP onset</td>
<td>14</td>
<td>9.8 ± 7.3</td>
<td>47</td>
<td>10.6 ± 5.7</td>
</tr>
<tr>
<td>MAC onset</td>
<td>14</td>
<td>19.2 ± 20.1</td>
<td>37</td>
<td>17.8 ± 18.0</td>
</tr>
<tr>
<td>Fundoplication</td>
<td>5</td>
<td>13.8 ± 28.3</td>
<td>24</td>
<td>11.2 ± 11.7</td>
</tr>
<tr>
<td>Home oximetry</td>
<td>22</td>
<td>70.5 ± 43.3</td>
<td>39</td>
<td>65.2 ± 45.8</td>
</tr>
<tr>
<td>Current age</td>
<td>22</td>
<td>9.6 ± 4.0</td>
<td>5</td>
<td>66.2 ± 114.2</td>
</tr>
</tbody>
</table>

*P value for comparison of groups 2 and 3.
Group 1, SMA 1 children not treated to prevent respiratory failure; group 2, SMA 1 children who underwent tracheotomy for respiratory management; group 3, SMA 1 children who used noninvasive mechanical ventilation and mechanically assisted coughing (MAC) instead of tracheostomy; BiPAP, bilevel positive airway pressure ventilation via nasal interface.
tracheotomy had been performed because of anoxic encephalopathy that had occurred during a cardio-
pulmonary arrest from airway mucus congestion before undergoing tracheotomy. FIFTEEN patients had indwelling tracheostomies before referral to us, from 7.6 ± 3.3 mos of age. Five others underwent tracheotomy before they were able to procure NIV and MAC (the delays were attributed to medical insurance delays). Five others underwent elective tracheotomy after having used NIV up to a mean age of 41.5 ± 18.3 mos (range 24–71 mos). Three of these patients were continuous bilevel PAP dependent and were having secretion aspiration to the extent that baseline SpO₂ dipped below 95%; two underwent tracheotomy because of frequent symptomatic bradycardias that resulted in oxyhemoglobin desaturation. Two other patients underwent tracheotomy at 12.0 ± 6.6 mos of age because of congenital bronchomalacia and anoxic encephalopathy, respectively.

Six patients had comprehensible speech at the time of tracheotomy and retained some ability to communicate verbally afterward. None of the 21 patients who had not developed the ability to speak before undergoing tracheotomy did so after tracheotomy. Considering the 47 group 3 (noninvasively managed) patients, six are over age 10 and 14 are over age 6. Pneumonia and respiratory failure were noted to have resulted from otherwise benign upper respiratory tract infections in at least 186 of 495 (37.5%) hospitalizations. In six cases, patients were admitted for gastrostomy tubes and intubated, failed extubation, and were transferred to our service for successful protocol extubation. The patients’ parents reported 605 avoided hospitalizations. In addition, there were episodes when MAC was successful in reversing oxyhemoglobin desaturations associated with airway mucus congestion attributable to difficulty controlling oral secretions on a daily basis.

Among the eight deceased NIV patients, three died suddenly at home from bradycardia to cardiac standstill. One died from an intracranial hemorrhage while intubated for respiratory failure. Two died from mucus plugs during upper respiratory tract infections when intubation was unsuccessful. One died from septic shock, and one died from a pulmonary embolism.

Of these 47 patients, 32 used high-span bilevel PAP only during sleep, six required it for more than 16 hrs/day and became hypercapnic and dyspneic when not using it, and nine required it continuously, with little or no breathing tolerance. All patients required high-span bilevel PAP continuously during upper respiratory tract infections. Of the 32 patients who used high-span bilevel PAP only during sleep, 10 could speak clearly, 16 had severe dysarthria, and six were averbal with no

<table>
<thead>
<tr>
<th>Age, yrs</th>
<th>No. of Hospitalizations</th>
<th>Patient-years</th>
<th>Hospitalizations per Patient-year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–3</td>
<td>25</td>
<td>67</td>
<td>0.37</td>
</tr>
<tr>
<td>3–22.5</td>
<td>12</td>
<td>89</td>
<td>0.13</td>
</tr>
<tr>
<td>Group 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–1</td>
<td>98</td>
<td>47</td>
<td>2.09</td>
</tr>
<tr>
<td>1–2</td>
<td>63</td>
<td>40.25</td>
<td>1.57</td>
</tr>
<tr>
<td>2–3</td>
<td>29</td>
<td>33.1</td>
<td>0.88</td>
</tr>
<tr>
<td>3–4</td>
<td>9</td>
<td>28.1</td>
<td>0.32</td>
</tr>
<tr>
<td>4–5</td>
<td>6</td>
<td>24.2</td>
<td>0.25</td>
</tr>
<tr>
<td>5–6</td>
<td>1</td>
<td>19</td>
<td>0.05</td>
</tr>
<tr>
<td>6–7</td>
<td>0</td>
<td>16</td>
<td>0.02</td>
</tr>
<tr>
<td>7–8</td>
<td>2</td>
<td>14</td>
<td>0.14</td>
</tr>
<tr>
<td>8–9</td>
<td>0</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>9–10</td>
<td>0</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td>10–11</td>
<td>0</td>
<td>6.5</td>
<td>0</td>
</tr>
<tr>
<td>11–12</td>
<td>0</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>12–13</td>
<td>0</td>
<td>0.8</td>
<td>0</td>
</tr>
<tr>
<td>Group 3 totals</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–3</td>
<td>190</td>
<td>120.35</td>
<td>1.58</td>
</tr>
<tr>
<td>3–5</td>
<td>15</td>
<td>52.3</td>
<td>0.29</td>
</tr>
<tr>
<td>5–13</td>
<td>3</td>
<td>84.3</td>
<td>0.04</td>
</tr>
<tr>
<td>3–13</td>
<td>18</td>
<td>136.6</td>
<td>0.13</td>
</tr>
</tbody>
</table>
ability to use air that crossed the vocal folds for communication. Of the six patients using high-span bilevel PAP 16–20 hrs/day, five have severe dysarthria and one is averbal. Of the nine who require continuous support, one can speak clearly, two have severe dysarthria, and six are averbal. Thus, 3 of the 47 NIV patients could communicate verbally, by contrast with the 6 of 27 patients in group 2 who were able to do so.

Of the 194 extubations done without subsequent tracheotomy, 6 of 62 (9.7%) were successful with conventional management, 28 of 58 (48.3%) were successful with conventional management plus bilevel PAP, and 61 of 74 (82.4%) were successful using the NIV protocol. In addition, 42 otherwise healthy children were intubated for brief (gastrostomy) surgical procedures. Twelve of 12 were successfully extubated using our protocol. Six of the other 30 failed conventional extubation and were transferred to our unit for successful protocol extubation to NIV.

**DISCUSSION**

This work provides anticipated outcomes for parents and clinicians deciding between three possible approaches to the typical SMA 1 patient. The approaches are palliative care avoiding intubation, tracheotomy, and the use of NIV and MAC both routinely and after extubation. In the first approach, despite letting nature run its course and avoiding ventilator use, most parents elect for gastrostomy to prolong effective nutrition, but these children usually die from respiratory failure before 12 mos of age.

Both NIV and tracheostomy can greatly prolong survival, thus far up to 14 and 22 yrs of age, respectively. Tracheostomy also significantly decreases hospitalization rates \( (P = 0.0001) \) until age 3, but it is associated with a statistically significant \( (P = 0.04) \) increase in hospitalization rates after age 5 compared with the NIV approach. Because most health care providers are familiar with it, parents are usually spared from having to provide continuous care for their children with tracheostomy tubes. Two of five children clearly died from complications related to their tracheostomy tubes, but, in addition, ventilatory support via tracheostomy has been associated with abnormal electrocardiograms, autonomic dysfunction, cardiac arrhythmias, trachitectasis, and sudden death from hemorrhage, mucus plugging, accidental disconnections, infection, and loss of autonomous breathing. On the other hand, at least three of the group 3 children who were too young to optimally cooperate with MAC and who died from airway congestion may have survived these episodes had they had tracheostomy tubes.

Immediately after tracheotomy, 25 of 27 patients permanently lost all ability to breathe unaided. This is in contrast with the group 3 patients who were extubated successfully to high-span bilevel PAP despite complete failure to ventilator wean (in 23 cases) and who subsequently weaned back to their preacute illness bilevel PAP regimens, often after returning home up to 3 wks later.

Twenty-one of 27 patients with tracheostomy tubes were averbal, including all of those who did not use NIV long enough to master speech before undergoing tracheotomy. On the other hand, only 13 of 47 NIV patients were averbal.

It can be argued that the patients with tracheostomy tubes might have been more severely affected because their first episodes of hospitalization for respiratory failure and gastrostomy placement occurred at significantly younger ages \( (P < 0.05) \) than for the NIV-managed patients. However, only half of the group 2 patients were placed on bilevel PAP, which was, in all cases, used at only low spans (for minimal assistance). This, too, could have resulted in earlier hospitalization for respiratory failure. There is also a tendency to perform gastrostomy during the same surgical intervention as tracheotomy. The group 1 patients were first hospitalized at significantly younger ages \( (P < 0.05) \) than the patients in groups 2 and 3. It is unclear whether this is because they were the most severely affected or whether not using bilevel PAP at all was responsible for the earlier hospitalizations. They also may not have benefited from the same degree of caution in avoiding respiratory tract pathogens or from the same nutritional interventions as the children from groups 2 and 3. Two of the group 1 patients did not undergo gastrostomy despite weight loss. Another limitation of this analysis is that it is retrospective. This makes the assignment of causality risky in that the groups may not be equivalent, because of subtle selection bias.

We have previously demonstrated that protocol extubations are significantly more likely to result in successful extubations than is conventional management \( (P = 0.006) \). This is the most important reason that these patients can survive in the long term without tracheostomy. Hospitalization rates diminish to levels equivalent to those of patients with tracheostomy tubes between ages 3 and 5, and then to significantly lower levels than for patients with tracheostomy tubes after age 5, when children can fully cooperate with MAC. The ability to fully cooperate with MAC also has been reported to be instrumental in avoiding hospitalizations for respiratory failure for children with muscular dystrophy. It is interesting to note that all of the successful nonprotocol extubations to bilevel PAP were performed on patients who used bilevel PAP nocturnally before acute respiratory failure; this may have favored successful extubation. These children, therefore, adapted more easily...
to it after extubation. Thus, extubation to full noninvasive ventilatory support can be critical for successfully extubating these children. Its routine nocturnal use from the time of diagnosis also reverses paradoxical breathing, prevents or reverses pectus excavatum, and promotes lung growth and chest development.2,5

The paradigms of not extubating patients until they are ventilator weaned, extubating without full (noninvasive) ventilatory support, using airway vibration methods rather than MAC to expulse secretions, and routinely administering oxygen rather than using the oximeter as feedback for effective assisted coughing must be rethought if long-term survival is to be achieved without unnecessary episodes of acute respiratory failure and tracheostomy.

Members of pediatric sections of national medical societies, including 75 intensivists, 61 physiatrists, and 51 neurologists and medical directors of neuromuscular disease clinics, responded to surveys regarding their recommendations when faced with an SMA 1 baby in respiratory distress. In general, it was found that physicians often intentionally withhold information on, or recommend against, potentially life-saving ventilator use, feeling that the quality of life of severely disabled children is too poor to warrant their survival and that their parents are too biased to make “appropriate” decisions.11,12 In a study of 70 health care professionals who were asked to judge the quality of life of children with SMA 1 using a Likert scale of 0–10, their replies averaged <2, and most said they would advise against any ventilator use.13 Yet, in a study of 104 care providers of 66 children with SMA 1, the responses were significantly more positive and were not significantly different than the responses of parents of unaffected children. With six typical SMA 1 children from group 3 now over age 10, it is quite possible that some will survive into adulthood. Thus, it is important for professionals not to impose their own concepts, values, and judgments onto people with disabilities14 and to be cognizant of their inability to gauge disabled patients’ life satisfaction and potential for social and vocational productivity. Unwarranted judgment of subjective factors associated with quality of life should not affect patient management decisions for the disabled.

REFERENCES
CME Article Number 1: J. Bach, et al.

1. By the age of 2 years, approximately 90% of children with spinal muscular atrophies (SMA) type 1:
A. Will be able to sit independently.
B. Will have developed respiratory failure or be unable to take food by mouth.
C. Will have minor delays in motor milestones.
D. Will have no evidence of disease.

2. In this study, the placement of a tracheostomy had what effect on the survival of children with SMA type 1:
A. Improved survival.
B. Worsened survival.
C. No effect.
D. Unknown effect.

3. In this study, the use of noninvasive ventilation had what effect on the survival of children with SMA type 1:
A. Improved survival.
B. Worsened survival.
C. No effect.
D. Unknown effect.

4. When compared with children that used noninvasive ventilation, those with a tracheostomy for respiratory management:
A. Displayed increased hospitalization rates before the age of 3.
B. Displayed decreased hospitalization rates after the age of 5.
C. Had initial hospitalization for respiratory failure at an older age.
D. Did not develop language and were nonverbal.

5. Pectus excavatum is:
A. An unavoidable manifestation of SMA type 1.
B. A surgically correctable disorder for SMA type 1 patients.
C. A result of chest wall muscle paralysis/weakness.
D. A condition shown to decrease respiratory function in SMA type 1 children.
The answers to any essay questions must be typed or computer printed on a separate piece of paper and attached to this page.

After finishing this exam:
1. Check your answers with the correct answers on page 379.
2. Complete the CME Evaluation and Certification on the following page and mail to Bradley R. Johns, Managing Editor, CME Dept.-AAP, American Journal of Physical Medicine & Rehabilitation, 7240 Fishback Hill Lane, Indianapolis, IN 46278.
3. This educational activity must be completed and postmarked by December 31, 2008. AAP Members may complete and submit this CME Answering Sheet and the following CME Evaluation and Certification page online through the members-only section of the AAP web page at www.physiatry.org.

Answering Sheet

AMERICAN JOURNAL OF PHYSICAL MEDICINE & REHABILITATION
Vol. 86, No. 5 • May 2007

CME Article Number 1: J. Bach, et al.
Circle the appropriate answers.

1. A  B  C  D
2. A  B  C  D
3. A  B  C  D
4. A  B  C  D
5. A  B  C  D
Was the article consistent with the stated objectives?  
1 2 3

Did reading this article prepare you to achieve its stated objectives?  
1 2 3

Is reading this article likely to enhance your professional effectiveness?  
1 2 3

Was the article format conducive to learning?  
1 2 3

Suggestions for future topics: ________________________________

I, ________________________________, certify that I have met the criteria for CME credit by studying the designated materials, by responding to the self-assessment questions, by reviewing those parts of the article dealing with any question(s) answered incorrectly, and by referring to the supplemental materials listed in the references.

This educational activity is designated for 1½ category 1 CME credits.

Indicate total credits claimed: __________ (maximum of 1½ credits)

Signature of Participant ________________________________ Date ______________________

Are you a member of the AAP?  □ Yes  □ No

Do you have an individual subscription to the Journal?  □ Yes  □ No

If you are not an AAP member or a Journal subscriber, have you enclosed payment of $15 with your exam?  □ Yes  □ No

Payment Options:  □ Check (Payable to AAP)  □ VISA  □ Mastercard

Card # ____________________ Exp. date __________ Amount $ __________

Signature ________________________________

Please provide your name as it should appear on your certificate of credit and also provide your mailing address.

Name ________________________________ Position ____________________

Institution _______________________________________________________

Street Address ___________________________________________________

City, State, Zip _________________________________________________

Phone ____________________ Fax ____________________ Email ______________

Office Use _______________________________________________________

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.
Medical Considerations of Long-Term Survival of Werdnig–Hoffmann Disease

ABSTRACT


Objectives: To report intercurrent nonrespiratory complications of unprecedented survival for Werdnig–Hoffman disease (spinal muscular atrophy type 1 [SMA 1]).

Design: A retrospective chart review and caregiver questionnaire for 103 consecutively referred SMA 1 patients for whom death was prevented during infancy.

Results: Overall, 15 of 63 (23.8%) respondents had severe, symptomatic bradycardias. Thirteen of 25 males had bilateral cryptorchidism, and two were unilateral. Ten of 42 (24%) respondents had recurrent oral candidiasis, and eight (19.0%) had recurrent nonoral candidiasis. All patients had hip dislocation/subluxation. They had collapsing spines and scoliosis by 1 yr of age. Ninety-six had indwelling gastrostomy/nasogastric tubes before 24 mos of age. Twenty-six underwent fundoplication to decrease reflux. Fifty-seven were on modified elemental diets with reported benefits. About one half had early pubarche. Three patients had episodes of acute pancreatitis. One untreated patient died of candidal endocarditis at 4 mos of age, and a second child had mitral valve candidiasis on autopsy.

Conclusion: Prolonged survival of SMA 1 results in a high incidence of concomitant medical conditions that need to be addressed.

Key Words: Mechanical Ventilation, Respiratory Paralysis, Quality of Life, Spinal Muscular Atrophy Type 1, Nutrition, Scoliosis, Autonomic Dysfunction, Cardiac Arrhythmias, Candidiasis
In a previous publication, Bach et al.\textsuperscript{1} have reported that respiratory muscle aids can be used to prolong the survival of children with spinal muscular atrophy (SMA) type 1. As a result, medical and surgical issues have arisen over time. The purpose of this work is to describe the incidence of and management rationale for complicating medical issues. There have been no previous long-term studies of SMA 1 patients.

**MATERIALS AND METHODS**

The statuses of all 106 SMA 1 patients who visited one neuromuscular disease clinic from March 1996 to April 2006 were reviewed. SMA 1 was diagnosed on the basis of DNA evidence of chromosome 5 exon 7 and 8 deletion in 95 of 103 children, affected siblings with genetically confirmed disease in five patients, and characteristic laboratory, muscle biopsy, and electromyography results in three children who did not undergo DNA testing. One child had an equivocal diagnostic workup; two were lost to follow-up after an initial visit before 12 mos of age; 11 were atypically, mildly affected; and the 18 untreated patients who died as infants (9.6 ± 4.0 mos of age) were eliminated from the survey.

This work was approved by our institutional review board. Data were gathered from clinical history, physical examination, and a follow-up questionnaire conducted by mail and telephone in June 2006 concerning cardiac and autonomic function, gastrointestinal function, and nutrition, infectious, developmental, orthopedic, and quality-of-life issues.

Inclusion criteria for typical SMA type 1 were inability to roll or sit unsupported at any time, paradoxical breathing, one or more episodes of respiratory failure requiring continuous respiratory assistance, loss of ability to receive any nutrition by mouth by 24 mos of age, and maximum vital capacity of 250 ml at any time. By 18 mos of age, all of the patients had little more than residual finger, toe, and facial movements. The 74 typical, treated SMA 1 patients were included. Twenty-seven were managed by ventilatory support via indwelling tracheostomy tube, and 47 were managed by using noninvasive ventilation (NIV) and mechanically assisted coughing (MAC).\textsuperscript{2,3}

The 74 patients’ medical histories were explored for gastroparesis, cryptorchidism, candidiasis, early pubarche, perceived benefits of using a modified elemental diet, hip dislocation, scoliosis, cardiac bradyarrhythmias, and positioning restrictions. The questions were described to the patient’s parents by a medical student. Symptomatic bradycardias were defined by loss of consciousness (LOC) associated with heart rates below 40/min in which oxyhemoglobin desaturation followed the bradycardia and there had been no concomitant acute respiratory difficulties. The parents were questioned concerning perceived benefits of a modified nondairy elemental diet consisting of formulas with high amino acid and polypeptide concentrations (Pediatric Vivonex and Tolerex; Novartis, Inc.). The diets were supplemented with B vitamins and minerals, glutamine, carnitine, long-chained free fatty acids, and lactobacillus.

**RESULTS**

Sixty-one of the 74 typical SMA 1 patients are currently 66.1 ± 44.8 mos of age, and 13 died at age 32.9 ± 50.4 mos of age. Of the five patients with tracheostomy tubes who are deceased, one died at 270 mos of age from a bradycardia after continuous respiratory support via tracheostomy tube from 2 mos of age. Three died from anoxia associated with purulent airway mucus congestion, and one died from persistent lung collapse attributable to congenital bronchomalacia. Of the eight deceased noninvasively managed patients, three died suddenly at home from bradycardia to cardiac standstill. One died from an intracranial hemorrhage, and two died from airway mucus plugging. Another died from septic shock, and one died from a pulmonary embolism.

**Cardiovascular and Autonomic Systems**

Two of 22 (9.1%) respondents with tracheostomy tubes and six of 41 (14.3%) without tracheostomy tubes had episodes of LOC associated with severe bradycardias. Three additional patients with tracheostomy tubes and four others without tubes had bradycardias to less than 40/min without LOC. Overall, 15 of 63 (23.8%) patients had severe bradycardias indicating apparent concomitant autonomic dysfunction.

**Gastrointestinal System**

Sixteen of the 27 patients with tracheostomy tubes had gastrostomy tubes placed before tracheotomy, eight had them placed concurrently, and three underwent gastrostomy after tracheotomy. All 27 underwent percutaneous endoscopic gastrostomy under general anesthesia. This resulted in six patients failing extubation. Of the 47 NIV-managed patients, 27 underwent placement of gastrostomy tubes under local anesthesia using the modified Stamm gastrostomy.\textsuperscript{4} 18 underwent percutaneous endoscopic gastrostomy under general anesthesia, in 17 cases along with Nissen fundoplication, and two have received all nutrition via nasogastric tubes from 4 mos of age to ages 12 and 8 yrs, respectively. In all, 26 of the 30 patients who underwent evaluation for gastroesophageal reflux also underwent Nissen fundoplication (under general anesthesia). Nine had fundoplication after gastrostomy, and one before gastrostomy. By avoiding general anesthesia or being extubated to continuous
NIV, none of these 47 patients had respiratory complications from these procedures.

Three patients had episodes of acute pancreatitis; two were felt to be secondary to gastroparesis, and the other patient was taking valproic acid in a research protocol.

**Nutrition**

Fifty-seven patients (20 with and 37 without tracheostomy tubes) were placed on modified elemental nutrition regimens via gastrostomy tubes from age 17.3 ± 19.1 mos for the patients using NIV and from 18.2 ± 16.5 mos of age for the patients with tracheostomy tubes. The seven regular-diet tracheostomized patients’ hospitalization rates were considered from 18.2 mos, and the eight regular-diet NIV users’ hospitalization rates were considered from 17.3 mos of age to match the ages of the modified-diet users. The modified-diet users had 0.09 ± 0.12 respiratory hospitalizations per year for 229 patient-years, and the 15 patients not on the diet had 0.07 ± 0.10 respiratory hospitalizations for 56 patient-years, with no statistically significant difference.

The parents of all the patients using the modified diet reported one or more of the following benefits: decreased episodes of perspiration and flushing, decreased airway secretions, more manageable secretions, more effective cough, and decreased abdominal distention. Nine of the 20 trached patients’ parents (45%) reported decreases in sweaty episodes andflushing, seven (35%) reported decreased airway secretions, nine (45%) reported greater facility in managing airway secretions, and two (11%) reported more effective expulsion of airway secretions using MAC. Among the parents of the 37 NIV users receiving the diet, 20 (54%) reported decreases in sweaty episodes and flushing, 28 (74%) reported decreased airway secretions, 27 (74%) reported greater facility in managing the secretions, and 11 out of 33 (33%) reported more effective secretion expulsion using MAC. There was decreased abdominal distention in 6 of 20 (30%) tracheostomized patients and in 21 of 35 (60%) NIV users. Some of the benefits also might have resulted from the institution of nocturnal NIV, which was begun within 4 wks of the modified diet for 15 of the 37 NIV-managed patients who used the modified diet.

**Genitourinary and Endocrine Systems**

In the tracheostomy-managed group, 56% percent (9 of 15) of the males whose testicles were examined were found to be cryptorchid bilaterally, and two additional males had unilateral cryptorchidism. In the NIV-managed patients, 4 of 10 males had bilateral cryptorchidism. One child underwent bilateral orchipexy at 7.5 yrs of age. Although it might be thought that the use of positive-pressure ventilation might facilitate testicular descent, this possible effect could not be ascertained from these data. In no cases were the males cryptorchidic at birth.

Fifty-eight percent (11 of 19) of tracheostomy-managed patients and 44% (15 of 34) of the NIV patients had pubic hair development at 53 ± 36 and 45 ± 29 mos, respectively. Four of 20 (20%) of the former and 3 of 34 (8.8%) of the latter group had underarm hair development at mean ages of 84 ± 49 and 44 ± 30 mos, respectively. Eight of 19 (42%) of tracheostomy patients and 13 of 33 NIV patients had adultlike hair distribution on the legs, back, or buttocks by 61 ± 43 and 40 ± 28 mos of age, respectively. Three of 10 (30%) of the tracheostomized boys and three of 16 of the NIV-user boys had increases in scrotum pigmentation (Tanner stage 2) by 88 ± 60 and 49 ± 46 mos of age, respectively. All of these means were in advance of the normal Tanner stage limits. Three of 10 (30%) tracheostomized girls and five of 18 (27.8%) NIV-user girls had telarche at 85.7 ± 3.8 and 91.9 ± 11.9 mos of age, respectively. One girl in the former group had menarche occur normally at 104 mos of age.

**Candidiasis**

Ten of 42 respondents (24%) had recurrent oral candidiasis, and 8 of 42 (19%) respondents had recurrent nonoral candidiasis. This included two of eight (25%) tracheostomy-user respondents and 8 of 34 (24%) NIV-user respondents who had oral candidiasis. Two of 8 (25%) of the former and 6 of 34 (18%) of the latter group had nonoral candidiasis at one or more of the following sites: tracheotomy/gastostomy/portacath sites, perineum, eyebrows, sides of the nose, neck, hands, fingers, armpits, chest, feet, toes, and scalp. One patient from the untreated group died of candidal endocarditis at 4 mos of age, and a second untreated child was found to have mitral valve candidiasis on autopsy. Neither patient was ever intubated, although both had indwelling gastrostomy tubes.

**Quality of Life—Positioning**

The mean time that the patients could sit at >45 degrees was 4.1 ± 4 hrs/day for 36 families questioned. A mean sitting angle of 52 ± 28 degrees was reported. Sitting was limited by difficulty managing airway secretions and dyspnea for 13 patients (four with and nine without tracheostomy tubes), dyspnea when not using NIV when seated for two patients, fatigue for five patients, musculoskeletal discomfort attributable to scoliosis (three patients), and coccygeal and buttck discomfort (three patients). Three others complained of boredom from decreased ability to use their activity-facilitating devices (Fig. 1). Six patients could not sit at all; one of
these patients had a tracheostomy tube, and five were NIV users.

Orthopedics

All 74 SMA 1 children seemed to have bilateral hip dislocation/subluxation on physical examination, and all developed increasing scoliosis. Although no typical SMA 1 patients have yet undergone surgical correction for scoliosis, four atypically mild SMA 1 patients underwent posterior arthrodesis for surgical correction of scoliosis despite long-term continuous dependence on NIV support. Two patients agreed to tolerate discontinuance of narcotic therapy to have nasogastric and translaryngeal tubes removed the day after surgery. They were extubated back to continuous NIV delivered via a 15-mm flexed mouthpiece diurnally and via a nasal interface during sleep. The other two were extubated on postoperative day 3 when narcotics and sedatives could be discontinued or minimized without excessive discomfort. One patient had a history of 11 hospitalizations for acute respiratory failure and had been extubated each time back to continuous NIV and aggressive MAC.

DISCUSSION

For parents who choose to have their children supported by the use of NIV and MAC or by tracheostomy, nonrespiratory complications can be anticipated.

Nutrition

It is clear that SMA patients can dramatically lose muscle strength during respiratory tract infections and other episodes of physiologic stress and undernutrition. In addition to metabolic aberrations associated with immobility, systemic illness, muscle denervation, and muscle atrophy, SMA patients have inborn metabolic abnormalities in mitochondrial fatty acid oxidation and carnitine metabolism, resulting in early fasting ketosis because of defective beta oxidation of fatty acids by muscle, similar to that seen in mitochondrial myopathies.4–7 Four hours or more of fasting results in ketoacidosis and respiratory muscle catabolism in these patients.5 Within 3 hrs of a normal meal, blood amino acid levels can decrease to levels that would not be reached until after more than 8 hrs of fasting in unaffected children. Also, by not efficiently metabolizing fatty acids, a major source of energy during fasting, abnormal depositions of fatty acid metabolites such as dicarboxylic acids occur in the blood and urine after any overnight fasting.5,6 Hypoglycemia can be precipitous and severe.

Thus, an elemental diet high in amino acids and polypeptides might benefit these children. Many parents have placed their children on diets high in carbohydrates, amino acids, and polypeptides, but low in fat. Such diets provide the muscles with usable energy substrates, thereby decreasing dependence on fatty acid oxidation and decreasing excessive accumulation of potentially toxic free fatty acids that can further damage muscle.6 This diet maintains more normal blood glucose levels during fasting, delays fasting ketoacidosis, and has been noted to normalize liver function enzyme levels. Provision of amino acids and short-chained peptides instead of complex dairy proteins facilitates glucogenesis and seems to have a beneficial effect on decreasing airway (and gastric) secretions and viscosity for some children. Further rationale for the possible benefits of a modified elemental diet have been summarized elsewhere.8

The hospitalization rates of children receiving elemental diets were not significantly different from those of nonmodified diet cohorts; however, only 15 of the latter could be studied. Despite this, 65% of the trachestomized users’ parents and 80% of the NIV users’ parents perceived multiple benefits. All of the benefits were significantly more frequent for the latter patients (P < 0.05) except for “decrease in sweaty episodes,” which was 9% more frequent in the trachestomized patients. This suggests a possible synergy between using NIV and the modified diet because the reported benefits of having fewer secretions, more manageable secretions, better coughing, and decreased abdominal distention may be more important for patients without tracheostomy tubes. The presence of tracheostomy tubes, causing secretions that necessitate airway suctioning an average of eight times a day,9 also facilitates their expulsion by invasive airway suctioning, and so the effect of the diet may be less apparent. Consistent claims of benefits in-

FIGURE 1 Nine-and-a-half-year-old girls with spinal muscular atrophy type 1, vital capacity 80 ml, lying on their sides and using an assistive device to permit upper-limb function. They were hospitalized and intubated at ages 1 and 4, but they depend on nasal ventilation up to 24 hrs/day.
dicate a need for further study regarding diet modification.

**Gastrointestinal Issues**

There was no increase in vomiting, regurgitation, or abdominal distension as a result of using NIV. However, the risk of gastroesophageal reflux is high in SMA 1 patients, especially in those having gastrostomy tubes, and 26 of 30 studied for reflux underwent fundoplication. The fact that nine patients underwent fundoplication after gastrostomy suggests the increase in reflux caused by alimentation via gastrostomy. Gastrostomy, however, can now be placed under local anesthesia, thereby avoiding respiratory complications associated with general anesthesia. Fundoplication, on the other hand, is not performed under local anesthesia. Because there was no apparent difference in hospitalization rates or mortality for patients as a function of having undergone this intervention, it may be best to promote gastrostomy under local anesthesia and reserve fundoplication for patients who are grossly symptomatic for reflux subsequently. Gastrostomy should be performed under local anesthesia and reserve fundoplication for patients who are grossly symptomatic for reflux subsequently. Gastrostomy should be performed under local anesthesia and reserve fundoplication for patients who are grossly symptomatic for reflux subsequently.

This is the first report of acute pancreatitis in children with SMA type 1. The incidence in our population was 3.6% (three cases, all with gastrostomy tubes), which is significantly higher than the overall incidence in American children of 0.2% (19/10,000). Two of the cases may be explained by gastroparesis, and one was associated with valproic acid use—both of which are known to predispose to pancreatitis. An indwelling gastrostomy tube also may be a causative factor for pancreatitis.

**Cardiovascular and Autonomic Systems**

There were statistically insignificant differences between the children with and without tracheostomy tubes regarding the number who had experienced bradycardia-associated LOC (<40/min). Overall, 23.8% of the SMA 1 patients had episodes of bradycardia-associated LOC. Although this study did not include children with SMA type 2, we have observed bradycardia-associated LOC in this population. Patients with episodic LOC were often evaluated for seizure activity but had normal electroencephalograms. The monitoring for bradycardias, along with the need for feedback for assisted coughing during upper-respiratory tract infections, are indications for home oximetry. Bradycardias seem to indicate autonomic instability, a complication already described for patients with adult-onset motor neuron disease. Others have reported cardiac arrhythmias and cardiomyopathies in SMA patients. If bradycardias are found to be triggered by preceding tachycardias, beta-blocker therapy can be tried. Otherwise, a cardiac pacemaker needs to be considered.

**Genitourinary and Endocrine Systems**

It is well known that boys with decreased abdominal pressure from abdominal muscle weakness tend to have cryptorchidism, as in prune belly syndrome. The rate of undescended testes is 3.7% (range, 3–5%) in normal newborn males, 1% at 3 mos, and 0.8% at 9 mos of age. Damage and volume loss of the testes is evident by 6 mos of age. The abnormal temperature of the cryptorchid testis increases the risk of carcinogenesis. Six to 10% of all testicular cancer originates in cryptorchid testes, with a cancer rate 1.75–3.5% over a normal lifetime, but this is otherwise rare in children, with about one case per 100,000 males. Because our oldest child with SMA 1 was 22.2 yrs old, the risk of an infant with SMA with untreated cryptorchidism being diagnosed with testicular cancer by 20 yrs of age is 0.1–0.2%.

This risk is certainly less than that of respiratory or other complications of orchidectomy under general anesthesia. Fortunately, the procedure can, and should, be done under local anesthesia. However, considering the very small risk of testicular cancer in these children, the fact that survival to adulthood must be considered unlikely, and the fact that the atrophied cryptorchid testicle can produce some testosterone for puberty, orchidectomy does not seem warranted.

Puberty normally occurs no earlier than at 84 mos in girls and 96 mos in boys. Pubarche and increased scrotum pigmentation began about 30 mos prematurely in SMA 1 patients. This suggests facilitation of adrenarche functions by failure of the normal survival motor neuron gene product. It should be pointed out that at least one atypically mild SMA 1 patient was reported to have delivered 2 healthy children by cesarean section.

**Candidiasis**

*Candida* usually affects only the oropharyngeal mucosa (thrush) and the perineum (diaper thrush). Newborn thrush is usually acquired during labor or during breast feeding. If thrush persists past 5 mos of age without ongoing antibiotic therapy or debilitation, further diagnostic workup is indicated. Causes of recurrent candidiasis in infants are malnourishment, low birth weight, and prematurity. The following infections do not occur in the immunocompetent patient: candidiasis of the esophagus, heart, urinary tract, gastrointestinal system, larynx, lung, blood stream, retina, and central nervous system.

Normal oral candidal flora has been reported in 48% of all infants, and oropharyngeal candidal infection occurs in up to 3% of the infant population (0–2 yrs). Approximately 17% of infants de-
velop candidal diaper rash.27 The annual rate for candidemia in patients with central venous catheters was reported as 3–4%.28 If a patient has candidemia, the risk of developing candidal endocarditis is approximately 6% and the rate of candidal endocarditis in patients with central venous catheters is up to 0.24%/year.28

Some risk factors for invasive candidiasis are prolonged antibiotic therapy, a central venous/arterial line, surgery, total parenteral alimentation, use of corticosteroids, endotracheal intubation, and prior candidal colonization.29 Patients with increases in *Candida* species in their stool of $\geq 10^4$ are at greater risk of invasive disease than those without such increases.30 Therefore, gastrointestinal tract colonization may be an important factor in the development of invasive candidiasis.30 In our patients, all of whom had indwelling gastrostomy tubes, local candidal infection may have been the primary cause of their invasive candidal diseases.

Endocarditis in children, usually associated with congenital heart disease, has an incidence of 0.05% in admissions to an active pediatric cardiac center.31 Fungal endocarditis makes up 5–10% of infective endocarditis, with the *Candida* species making up 67% of fungal endocarditis infections.32 The number of reported cases of fungal endocarditis in children has increased, largely among neonates receiving intensive care.33 Heart surgery, rheumatic valvulitis, prosthetic heart valves, long-term intravenous/umbilical catheters, tracheal traumatization, or spread from bronchopulmonary infection can all give fungi access to heart valves.32 This is the first report of children with SMA type 1 with recurrent candidiasis or candidal infection of heart valves. More than one quarter (26.2%) of these patients had nonmucocutaneous candidiasis, including two with candidal heart valve encrustations. Nonmucocutaneous candidiasis (especially when recurrent) seems to only occur in immunocompromised individuals. Some caregivers have suggested that treatment with *Lactobacillus acidophilus* seems to have decreased recurrent candidiasis rates.

Orthopedic Issues

Minimization of limb contractures and scoliosis, and the effects of immobility and inability to bear weight on the skeletal system, need to be considered. All children with SMA 1 have bilateral dislocated/subluxed hips. Because all three of our SMA 2 patients who underwent surgical procedures to better form the hip joint went on to develop chronic hip pain as adults, and none of our 140 SMA 1 and 2 patients who did not have hip surgery have thus far gone on to develop hip discomfort, including 18 SMA 2 patients as adults, it is apparent that no hip dislocation-reduction surgery is indicated for these nonambulatory patients. Tilting to semistanding position with support, calcium and vitamin D supplementation, and, possibly, alendronate sodium may hinder osteoporosis; however, the latter has not been approved for use in children.

In a study of patients with pediatric neuromuscular diseases, it was determined that patients perceive limb contractures to impede activities of daily living and to cause considerable discomfort, especially when being dressed.34 Thus, besides the use of thrice-daily limb-articulation mobilization, musculoskeletal release surgery should be considered for the upper limbs as well as the lower limbs. Fortunately, children with SMA 1 tend to have fewer fixed contractures than children with SMA 2 because of their more severe muscle weakness and less strength imbalance across limb articulations. Unfortunately, the risk of cardiorespiratory complications of surgery is likely to be greater in SMA 1 than in SMA 2. Most recently, the use of NIV has been reported to facilitate safe general anesthesia for orthopedic surgery and to prevent postoperative respiratory complications and the need to consider tracheotomy for NIV-dependent children.35

There are, as yet, no reports of spinal instrumentation to ameliorate scoliosis for patients with typical SMA type 1. We have had four atypically mild SMA type 1 patients and 11 with SMA type 2 undergo spinal instrumentation in their teen years, with excellent correction and absence of respiratory complications, even though six were extubated to NIV despite being unable to ventilator wean.35 In all cases, the patients were thoroughly trained in NIV before surgery and were extubated to full, noninvasive ventilatory support. Without specific knowledge of the extubation protocol to NIV, these patients would have had to undergo tracheotomy.2,3

Positioning

Although one usually is more functional when sitting than when reclining, because of the upper-extremity and finger function afforded by supporting the elbow and forearm using an upper-extremity dynamic orthosis,36 these patients tend to prefer to remain side-lying (Fig. 1).

REFERENCES

5. Crawford TO, Sladky JT, Hurko O, Besner-Johnston A,
May 2007

Considerations in Werdnig–Hoffmann Disease


The Impact of the Inpatient Rehabilitation Facility Prospective Payment System on Stroke Program Outcomes

ABSTRACT

Objective: To examine the impact of the Medicare inpatient rehabilitation facility (IRF) prospective payment system (PPS) on outcomes in a stroke rehabilitation program.

Design: An analysis was performed on a database including 945 stroke patients admitted to an inpatient stroke rehabilitation program 5 yrs before implementation of the IRF PPS and 3.5 yrs after implementation. Patients were classified with regard to stroke location (left vs. right), level of cognitive impairment, presence/absence of unilateral neglect, and level of depressive symptomatology. Functional status was evaluated at time of admission and discharge by functional independence measure (FIM). Other outcome measures included length of stay (LOS) and discharge destination. The impact of IRF PPS on LOS, progress in rehabilitation, and discharge destination was examined via univariate analyses of covariance and logistic regression.

Results: Patients admitted after implementation of the IRF PPS had shorter LOS but made less progress, had lower functional levels at discharge, and had higher rates of institutional discharge.

Conclusions: Although associated with decreased LOS, implementation of the IRF PPS was associated with decreased functional gains, lower discharge FIM levels, and higher rates of institutional discharge. Cost savings associated with the PPS must be considered in light of these untoward outcomes.

Key Words: Stroke, IRF PPS, Rehabilitation Outcomes
Implementation of the Medicare inpatient rehabilitation facility (IRF) prospective payment system (PPS) in 2002 has generated concern regarding the potential impact of the new system on rehabilitation outcomes. In this system, reimbursement for a rehabilitation admission is fixed, determined by specific diagnosis and by certain modifiers, such as the patient’s age, level of initial functional impairment, and other medical complications. The IRF PPS is designed to provide hospitals with a powerful financial incentive to minimize lengths of stay (LOS). By providing a fixed prospective payment for each patient, hospitals that discharge patients more rapidly can admit more patients (i.e., turn over beds more quickly) and, as a result, increase their revenue. However, with shorter LOS comes the distinct risk that patients will be discharged before essential rehabilitation goals can be met, or at functional levels that increase the level of burden placed on caregivers.

Inpatient rehabilitation LOS had already declined significantly before IFR PPS. A recent study involving 744 facilities and 148,807 patients conducted during a 7-yr period before implementation of the IRF PPS has revealed an 8-day decrease in median LOS with functional status at time of discharge and percentage of patients discharged home remaining stable. Although these findings suggest that LOS could be decreased without affecting a patient’s likelihood of being discharged home, the authors report that the mortality rate was higher for these patients in the 6 mos after discharge, raising the possibility that important medical issues had not been addressed adequately.

No studies to date have evaluated the effect of the IRF PPS on specific diagnostic groups such as stroke patients. Stroke constitutes one of the largest diagnostic groups treated within rehabilitation facilities and one of the most variable in terms of case complexity. IRF PPS attempts to adjust the amount of reimbursement in accordance with level of functional impairment and presence of certain comorbidities (which, in principle, should allow longer rehabilitation stays for more severely impaired patients). It is not clear, however, whether this system of adjustment is adequate or, in particular, whether the currently used weighting system accurately predicts case complexity. We recently have demonstrated in a pre–IRF PPS stroke population that the presence of unilateral spatial neglect resulted in longer LOS and slower improvement, even when adjusting for initial level of functional impairment. Yet unilateral spatial neglect is not used as a determinant of case complexity in the IRF PPS system.

The current study represents a preliminary investigation of the effect of the IRF PPS on stroke functional outcomes within an acute rehabilitation hospital. Data related to stroke admissions, including admission and discharge functional independence measure (FIM), discharge setting, and cognitive and psychological indicators, were collected starting in 1997 as part of an ongoing research program. Five years of admissions before implementation of the IRF PPS, and 3.5 yrs of admissions subsequent to initiation of the system, were available for analysis.

We compared stroke patients admitted before IRF PPS implementation with patients admitted after initiation of IRF PPS to determine whether post–IRF PPS patients had differences in LOS, functional improvement, and functional levels on discharge. Most importantly, we evaluated whether after implementation of IRF PPS there were differences in the proportion of patients discharged home or to the community, with a corresponding increase in admissions to institutional settings such as skilled nursing facilities.

Regarding specific subsets of stroke patients, we evaluated whether right hemisphere stroke patients, who demonstrate poorer rehabilitation outcomes, including longer lengths of admission and greater dependence on others for performance of activities of daily living, were disproportionately affected by IRF PPS in terms of functional gains and discharge functional levels. Within this subset of stroke patients, we anticipated that patients with unilateral spatial neglect, who progress even more slowly and have longer LOS than other right hemisphere stroke patients, would be most severely affected, particularly in terms of an increased likelihood of discharge to institutional settings such as skilled nursing facilities.

**METHOD**

**Participants**

Data were obtained from 945 patients (466 men; 479 women) admitted to an inpatient stroke rehabilitation program during 8.5 yrs. Ninety-six percent of the sample were white, and 2% were African American. Their mean age was 72.66 yrs (SD = 11.63). Forty-nine percent of the patients had right hemisphere stroke, 41.5% had left hemisphere stroke, and 9.5% had bilateral stroke. Twenty-one percent of the sample had a previous stroke.

**Procedure**

All patients with a diagnosis of stroke who were able to complete a neuropsychological screening as part of the cerebrovascular accident (CVA) program protocol including the Cognistat and the Geriatric Depression Scale (GDS) were included in the current sample. A subset of patients also were administered a structured letter-cancellation
Patients unable to complete the screening because of aphasia, profound cognitive deficits, or perceptual deficits were excluded from the sample. In addition, patients whose stroke had occurred more than 90 days before screening were excluded. Overall, from a total of 1281 referrals, 945 patients were included in the sample; 536 patients were admitted during the period before IRF PPS implementation, and 409 patients were admitted after its implementation.

Stroke patients were administered the screening tasks approximately 4 days after rehabilitation admission. All tests were administered by a clinical neuropsychologist or a trained testing technician working under the neuropsychologist’s supervision. The Cognistat was administered according to the instructions provided by Weintraub and Mesulam.10

The FIM12 was administered on admission and discharge by members of the patient’s rehabilitation team. All FIM raters were trained and accredited in accordance with Uniform Data System standards.

Measures
Letter-Cancellation Test

This test consists of 60 target letters A randomly dispersed among 314 nontarget letters on a structured stimulus array. Thirty of the target letters are on the left side of the stimulus array, and the other 30 are on the right side. Subjects were required to cancel out all letters A on the page. The number of targets omitted was scored with a maximum score of 60. Classification of unilateral spatial neglect was determined via criteria developed in previous research13,14 and required: (a) more than eight errors of omission, and (b) an asymmetry score (derived by dividing targets correctly cancelled on the left by total number of targets cancelled) of <0.50, reflecting a left side bias with regard to number of errors.

Cognistat

The Cognistat8 is a test of cognitive functioning that assesses multiple skill areas, including orientation, attention, memory, language, construction, calculations, similarities, and reasoning. A total score for each individual skill area and for the number of skill areas exceeding the established cutoff was computed.

GDS

The GDS9 is a 30-item questionnaire designed specifically for detection of depression in elderly populations. Internal consistency was 0.90.

FIM

The FIM12 is an 18-item indicator of functional status devised to determine the level of assistance required by a person with disability to perform basic life activities safely and effectively. It includes items assessing self-care, sphincter control, transfers, locomotion, communication, and aspects of cognitive functioning. Internal consistency was 0.93 and 0.95 for admission and discharge FIM, respectively.

RESULTS
Descriptive Statistics

Participants entered the rehabilitation hospital with an average FIM score of 62.68 (SD = 17.74). They left the hospital an average of 21.35 days later (SD = 10.58) with a mean FIM score of 90.58 (SD = 20.33). On average, FIM scores improved significantly from admission to discharge ($t_{943} = -61.38$, $P < 0.001$). Patients’ mean score on the GDS was 7.51 (SD = 6.51).

Comparison of pre and post Implementation of IRF PPS

Table 1 provides a comparison of pre– and post–IRF PPS patients on demographic and admission variables and also on the main outcome measures: discharge FIM score, change in FIM score from admission to discharge, and LOS. Patients who entered the rehabilitation hospital after implementation of IRF PPS had more years of education, a shorter period of time from stroke to assessment, higher GDS scores, greater cognitive impairment, lower discharge FIM scores, less change in their FIM scores, and shorter LOS than those who were in the hospital before IRF PPS implementation. Table 2 presents the results of the $\chi^2$ tests comparing patients before and after IRF PPS implementation on categorical study variables, including several demographic factors, stroke-related variables, and the discharge setting variable. Patients were more often discharged to an institutional setting after implementation of IRF PPS than before IRF PPS implementation.

In univariate analyses of covariance, controlling for years of education, time from stroke to assessment, GDS score, and cognitive impairment, post–IRF PPS patients continued to have lower discharge FIM scores [$F(1, 1920) = 8.37$, $P < 0.01$], shorter LOS [$F(1, 1920) = 41.84$, $P < 0.001$], and less change in FIM scores [$F(1, 1920) = 39.70$, $P < 0.001$].
than pre-IRF PPS patients. In addition, a logistic regression controlling for years of education, time from stroke to assessment, GDS score, and cognitive impairment revealed that patients who were in the hospital after implementation of IRF PPS were more often discharged to an institutional setting than their counterparts who were in the hospital in the years before IRF PPS implementation (Wald criterion z = 16.59, P = 0.001).

In addition to the average pre–post IRF PPS differences described above, three figures were created to illustrate year-by-year changes in discharge FIM, LOS, and number of institutional discharges (Figs. 1–3). As can be seen from these figures, fairly abrupt changes occurred when IRF PPS was implemented in 2002: discharge FIM and LOS decreased, whereas the number of institutional discharges increased.

**Group Differences**

We next determined whether the implementation of IRF PPS differentially affected certain groups of stroke patients. We eliminated from the analyses patients who had a bilateral CVA and repeated the above analyses, including location of stroke as an independent variable to examine whether the difference between pre- and post-IRF PPS patients varied depending on location of the stroke. Results were identical to the larger sample, with none of the IRF PPS × stroke location interactions being significant.

To determine whether the effect of IRF PPS implementation varied depending on admission FIM score, we conducted a median split of the entire sample. Those with admission FIM scores of 64 or higher were placed in the high group, and those below 64 were placed in the low group. This categorical variable was included as an independent variable in the univariate analyses of covariance and as an interaction variable in the logistic

| TABLE 2 | \( \chi^2 \) tests comparing pre– and post–prospective payment system (PPS) patients on categorical study variables |
| --- | --- | --- | --- |
| Gender | Pre–IRF PPS | Post–IRF PPS | \( \chi^2 \) Value |
| Men | 261 | 205 | 0.19 |
| Women | 275 | 204 | |
| Marital status | | | |
| Single | 40 | 28 | 0.61 |
| Married | 294 | 217 | |
| Widowed | 162 | 131 | |
| Separated/divorced | 40 | 33 | |
| Type of CVA | | | |
| Hemorrhagic | 55 | 43 | 0.02 |
| Nonhemorrhagic | 480 | 365 | |
| Location of CVA | | | |
| Left | 214 | 178 | 2.40 |
| Right | 274 | 188 | |
| Bilateral | 48 | 42 | |
| Previous CVA | | | |
| Yes | 114 | 84 | 0.07 |
| No | 421 | 324 | |
| Discharge setting | | | |
| Home | 486 | 323 | 24.80*** |
| Other institutional setting | 45 | 79 | |

IRF, inpatient rehabilitation facility; CVA, cerebrovascular accident.

*** \( P < 0.001 \).
regression. The above findings were retained regardless of admission FIM level.

To determine whether the effect of IRF PPS implementation varied depending on whether a patient met criteria for left neglect, we examined only those patients with right hemisphere stroke who were able to complete a regulation letter-cancellation task. Results for this sample were the same as for the larger sample, with none of the IRF PPS × neglect interactions being significant.

**DISCUSSION**

We found that since the introduction of the Medicare PPS, LOS among stroke patients had decreased by 4.5 days. In addition, discharge FIM scores were lower than before PPS implementation, and patients showed less functional gain during the admission. There also was an increase in the number of patients discharged to institutional settings. Whereas before PPS only 9% of stroke...
patients were discharged from this acute rehabilitation hospital to institutional settings, the rate rose to 19% after introduction of the PPS.

These findings capture the negative outcomes that had been feared when the new payment system was proposed: that the financial incentive to reduce LOS would result in patients being discharged earlier at lower functional levels, having made fewer functional gains. That changes in rehabilitation practice would occur as a result of the new system is not surprising in view of a recent estimate that PPS payments are 37% lower than actual rehabilitation costs. This would seem to leave rehabilitation facilities with two options: either discharge patients sooner, or decrease the intensity of services provided.

It could be argued that these patients may simply continue their recovery through use of outpatient rehabilitation services in the community and eventually achieve levels of functioning comparable with those displayed by patients at the time of discharge before implementation of PPS. To date, there are no data to support this interpretation. On the basis of trends associated with decreased LOS before IRF PPS implementation, Ottenbacher et al.2 found that whereas level of discharge functioning remained stable during this period of LOS decline, the mortality rate for discharge patients increased.

Of greatest concern is the increased rate of institutional discharges after implementation of the PPS. Whereas before the PPS, longer LOS were associated with increased Medicare expenditure, this investment in resources increased the likelihood of a home discharge. With shorter LOS and increased rates of nursing home discharge, the Medicare cost has declined, with the financial burden shifted to another source. Any calculation of the savings associated with IRF PPS should take into account the costs associated with higher rates of institutionalization, even if it does not come from the same health care fund.

DeJong et al.5,16 recent study of stroke patients involving three rehabilitation facilities found no change in LOS after implementation of the PPS, with modest declines in FIM discharge scores and amount of functional gain but no change in the rate of institutional discharges. The apparent discrepancy between the current findings and DeJong et al.5,16 findings might be attributable to the time frame studied. Although the exact time frame studied by DeJong et al.16 is not specified, it seems to involve the 1- to 2-yr period immediately after implementation of the PPS. The current study has included a 3.5-yr period after PPS implementation. As DeJong et al.16 note, the time period in their investigation may have been too brief to capture how facilities adjusted to the new system.

Contrary to our predictions, we did not find that PPS differentially affected right and left hemisphere stroke patients, nor did we find a differential effect for patients admitted at a low vs. high FIM level. However, among patients with neglect, the percentage who required institutional placement rose from 14% before implementation of PPS to 32% after implementation. Although this difference did not achieve statistical significance, most
likely because of the relatively small number of neglect patients, this trend warrants further study. Factors other than the implementation of PPS may have contributed to the decreased LOS and poorer functional outcomes in patients admitted after 2002. As evident in Figure 2, despite some year-to-year variability, there was a trend toward shorter LOS even before the introduction of PPS. Immediately after implementation, however, there was a steep decline in LOS beyond what could be attributed to a simple continuation of this trend.

One factor possibly contributing to the reduced FIM scores at discharge is that patients were admitted at lower functional levels and represented a more impaired sample of stroke patients. As measured by FIM, however, patients before and after implementation were admitted at similar levels of functional impairment. If these admissions were in fact more impaired, longer, rather than shorter, LOS would be anticipated.

For reasons that are not clear, patients admitted after implementation of PPS had higher levels of depressive symptomatology. Previous research has suggested that depressive symptoms may be associated with slower progress in rehabilitation. However, the differences we obtained controlled for depressive symptoms. Moreover, elevated GDS scores have been associated with longer rather than shorter LOS.

Stroke patients admitted since PPS implementation had higher levels of cognitive impairment. In view of the central role of learning/education in rehabilitation, a conceptual link between cognitive dysfunction (e.g., deficits in information-processing capacity or learning ability) and reductions in functional outcomes in rehabilitation is plausible. Along these lines, cognitive impairment measured by the Cognistat has been associated with poorer functional rehabilitation outcomes in stroke patients. However, our previous research involving a subset of the current stroke population has failed to find a relationship between Cognistat performance and FIM outcomes. In addition, we controlled for cognitive impairment in our analyses. In any case, it is difficult to conceive of a scenario in which greater cognitive impairment would result in shorter rehabilitation hospital LOS.

The current study has several limitations. First, because participants were drawn from a single facility, the generalizability of the current findings to the broader population of stroke rehabilitation patients might reasonably be questioned. Also, the representativeness of the sample may be questioned, particularly given the low proportion of minorities. Nonetheless, the scarcity of data on neglect patients might reasonably be questioned. Additionally, the representativeness of the sample may be questioned, particularly given the low proportion of minorities. Nonetheless, the scarcity of data on neglect patients might reasonably be questioned.

Second, we did not conduct follow-up assessments to determine whether there were long-term negative effects related to the earlier discharges. Over time, post-PPS patients may have achieved levels of functioning comparable with those of patients who had the benefit of a longer rehabilitation admission and who were discharged at higher functional levels. Similarly, we do not know whether some of the patients discharged to nursing homes as opposed to home eventually returned to the community. More importantly, we do not know whether the shorter admissions with discharge at lower functional levels were associated with adverse medical consequences (e.g., higher rates of hospital admissions) or, as has been suggested by Ottenbacher et al., with higher mortality rates.

Finally, in attempting to capture changes occurring during an 8.5-yr period of time via a clinical sample, there are numerous confounding factors that could have inadvertently influenced our findings. For example, with implementation of the PPS system and reliance on the admission FIM score as a means of documenting burden of care, pressures to score downward would seem inevitable because lower scores result in increased reimbursement. This type of variation in scoring from before to after PPS implementation could pose serious problems for interpreting differences in FIM scores over the two time periods. Although we cannot rule out variations in FIM scoring as a confounding factor, we note that admission FIM scores in the current sample did not decrease after implementation of PPS as might be anticipated, but actually increased slightly. In light of this finding, it seems unlikely that the post-PPS sample would be so much more impaired than the pre-PPS sample and result in such a large increase in institutional discharges. Nonetheless, the point remains that numerous factors other than introduction of the PPS may have influenced the changes we observed in functional outcomes and discharge destinations.

Although the precise mechanism through which pre- vs. post-PPS implementation change occurred is unclear, the current findings indicate that the primary goal of PPS, to reduce costs associated with rehabilitation by decreasing LOS, was achieved in the study facility. Although not conclusive, the current study raises the possibility that this goal may have been achieved at the cost of patients being discharged at lower functional levels (therefore requiring higher levels of assistance from caregivers) and by higher rates of institutional discharge. Further research is necessary to determine the relative contribution of the PPS vs. other as-yet unidentified factors.
CONCLUSIONS

Although LOS among stroke patients decreased after implementation of the IRF PPS, patients showed less functional gain during the admission, were discharged at lower FIM levels, and had higher rates of institutional discharge. Although the current findings do not establish a causal relationship between implementation of the PPS and these negative outcomes, it does suggest the need for additional research. If a causal link is established, cost savings associated with the PPS must be considered in light of these untoward outcomes.

ACKNOWLEDGMENTS

We gratefully thank Quintin Poore, PhD and Peter Gernert-Dott, PhD of Sunnyview Hospital and Rehabilitation Center for their assistance with data collection.

REFERENCES

The Clinician Effect on “Objective” Technical Components of the Electrodiagnostic Consultation

ABSTRACT


Objective: To examine the impact of clinician factors on technical data within an electrodiagnostic consultation for low-back pain and spinal stenosis.

Design: Examiner differences on single-segment paraspinal mapping scores and other findings were examined in a prospective, masked, double-controlled trial involving 150 people aged 55–80 yrs who were selected for no symptoms, back pain, or possible spinal stenosis.

Results: Unmasked clinicians were more variable than masked physicians ($F_{2.219} = 4.808, P = <0.01$) and gave lower scores to people they felt had mechanical back pain. The percentage of inadequate segmental scores differed among clinicians (0–16.6%, $F_{8.226} = 4.170, P < 0.001$), with fellows having more difficulty than faculty (11.76 ± 32.38% vs. 0.75 ± 8.67%) ($t_{233} = 3.753, P < 0.001$). Correction of clinician bias improved the relationship between paraspinal score and subjects’ ability to walk (weighted regression $R^2 = 0.129, B = -0.047, P < 0.001$; unweighted regression $R^2 = 0.090, B = -0.045, P < 0.001$).

Conclusions: Objective testing is adversely affected by clinician factors including prejudgment, experience, and individual idiosyncrasies. Less variation is found in more codified procedures. For electrodiagnostic consultation, correction of variability improves the relationship of test results to disability.

Key Words: Electromyography, Reliability, Bias, Masking, Back Pain
Most laboratory tests have known accuracy, validity, and reliability. Because they are technical in nature, no examiner bias is assumed. But medical diagnoses are often based on the conclusions of a consultant who combines objective tests with clinical judgment. Examples of tests that are actually consultation involving some objective measures include arthroscopy, colonoscopy, slit lamp examination, and radiologic and pathologic examinations. Choice of test components, skill in performing tests, and integration of test findings into a final diagnosis are all related to some extent to clinical judgment, whereas some test components themselves are specific codified procedures that should be objective.

Electromyography is often considered a test, but according to the American Association for Neuro-muscular and Electrodiagnostic Medicine, it is a complex consultation. The electrodiagnostic consultation (EDX) for spinal disorders involves objective testing intertwined with subjective judgment. History and physical examination lead to a differential diagnosis and the selection of different nerves and muscles to test. The tests themselves involve technical measures such as the conduction velocity of a nerve, the presence of fibrillations in a muscle, and the percentage of polyphasic motor units in a muscle. Test results and clinical suspicion may lead to unplanned explorations. The final diagnosis depends on clinical judgment as well as technical data.

EDX is, thus, a good model for examining whether subjective clinical judgment has an impact on more objective testing. The impact of subjectivity on variation in individual test results has not been explored, and the clinical consequences of subjectivity on final diagnoses are not known.

The general hypothesis of the current paper is that examiner subjectivity is a clinically relevant, sometimes detrimental, but modifiable factor in the technical portion of EDX. A number of specific hypotheses are explored, including:

1. Subjectivity is related to clinical presentation. Specifically, an unmasked examiner's EDX scores vary more than a masked examiner's scores in relation to the clinical diagnosis.
2. Examiners differ significantly in their scoring of people with the same spinal disorder. On average, people with the same problem should have the same amount of denervation. Statistically significant differences between the examiners are attributable to the examiners themselves.
3. Subjectivity can be minimized by more codified protocols. Specifically, the interexaminer variation with highly codified paraspinal mapping is less than that the variation in limb EDX findings.
4. Skill impacts the validity of results. Specifically, fellows report more invalid scores on paraspinal mapping testing than do faculty.
5. Adjusting for subjectivity improves results. Specifically, mathematical adjustment for individual clinical bias improves the relationship between test results and clinical measures.

**METHODS**

The Michigan Spinal Stenosis Study's overall methodology has been described in detail elsewhere. In an institutional review board–approved trial, subjects filled out questionnaires, performed ambulation tests, and underwent a codified but unconstrained spine history and physical examination, resulting in a clinical impression of spinal stenosis, mechanical back pain, or no symptoms. They underwent EDX testing and magnetic resonance imaging. Subjects who continued to qualify for inclusion were invited for a repeat protocol after 18 mos.

Subjects included 150 people ages 55–80 who were prescreened for risk of neuropathy and contraindications. They comprised three groups—community volunteers with no spine symptoms, and people recruited from clinically ordered magnetic resonance scans who, on review, seemed to have radiologic spinal stenosis or no stenosis and (on chart review) no leg pain. Although these preliminary groupings assisted in subject recruitment, the final clinical diagnosis used in the current study was made by a physiatrist who reviewed extensive patient history information and performed a verbal history and a comprehensive spinal physical examination but did not see magnetic resonance imaging or EDX information.

One hundred forty-eight of the 150 initial subjects and 87 of the 88 subjects returned for follow-up and underwent the EDX testing, for a total of 235 cases available for analyses. The three exclusions were attributable to inadequate EDX testing.

A masked EDX specialist performed a detailed electrodiagnostic study. To acclimatize the subject as part of the masking procedure, and as a measure of interrater variability, the unmasked physiatrist performed a very limited electromyogram of a single paraspinal level randomly chosen ahead of time by an assistant. Because masking of electrodiagnosticians had not previously been validated, we asked questions about whether any potential unmasking occurred, and whether clues such as pain behavior, arthritis, or muscle atrophy were related to the actual diagnosis. Inappropriate clues, almost all verbal hints, were detected in only 18 subjects. Half of these actually pointed the investigator in a direction different from the final EDX diagnosis,
and none hinted at a difference between stenosis and back pain.6

Electrodiagnostics were performed with a Nicolet Viking II (Nicolet Biomedical, Madison, WI) using standard techniques.4,5 In five limb muscles, fibrillations were scored 0–4+, and the morphology of 10 motor units was informally assessed.7,8 Nerve conduction studies were also done, with the temperature controlled at \( \pm 32^\circ \text{C} \).

The MiniPM version of paraspinal mapping, an anatomically validated, quantified scoring system for needle electromyography of the paraspinal muscles, was used.2,9–13 The ranges of normal and interrater reliability \((r = 0.830, P = 0.041)\) have been established. The initial description was published in the Archives of Physical Medicine and Rehabilitation,2 but an up-to-date description is maintained through a publication-quality course handout from the American Association of Neuromuscular and Electrodiagnostic Medicine.3

Briefly, paraspinal mapping includes palpation of the inferior border of the three lowest lumbar spinous processes and the midpoint between the posterior superior iliac spines, measuring 2.5 cm laterally, and for the L3, L4, and L5 spinous processes) 1 cm cranially. From each of these four locations, a 50- to 75-mm monopolar electromyographic (EMG) needle is inserted at a 45- to 60-degree angle to the surface in three different directions—cranially 45 degrees, directly across to the spinous process, and caudally 45 degrees—and advanced through the muscle in 5-mm movements to detect abnormal muscle-membrane instability. Any instability found (positive waves or fibrillations) must last more than 1 sec and be reproducible. Scores for the most medial centimeter are scored separately from the more lateral components of the insertion. Depending on the severity of findings, scores can range from 0 to 4 in any of 24 total locations. A total score for the side (number of positives) is determined.

Nine electrodiagnostician/physiatrist examiners variously performed the role of clinical examiner and electrodiagnostician. They had trained in eight different residency programs. Three were faculty physicians, board certified in physical medicine and rehabilitation, electrodiagnostic medicine, and pain medicine. Six had completed their physical medicine and rehabilitation residency and were in a fellowship designed to qualify them for both other boards.

Testing of Individual Hypotheses

The unmasked examiner placed the subjects into three groups—clinical spinal stenosis, mechanical back pain, and asymptomatic. The unmasked examiner’s paraspinal mapping score was compared with that of the masked electrodiagnostician at the same location. The total paraspinal mapping scores of each clinician within each diagnostic group were summed and compared among the different clinicians.

Analysis of variance was performed on paraspinal mapping scores, mean and maximum values of fibrillation scores, and motor unit scores of the masked clinicians.

We quantified the number of inadequate insertion sites, levels, and examinations by examiner, looking for individual differences and differences between experienced and relatively inexperienced clinicians. In paraspinal mapping, a score of 0, indicating inadequate testing, is obtained at a particular location when more than two motor units are on the screen (poor relaxation) or when less than 1 cm of adequate probing is done because of contact with bone or adipose. An entire level (e.g., left L4) is considered unreliable if zeroes are present in all six scoring locations at that level. A complete paraspinal examination on one side of a subject is considered inadequate if two or more levels on a side are inadequate.3

A subjectivity ratio—the average score of an unmasked clinician, divided by the average score of all of the other masked clinicians who tested the same subject in the same location—was calculated. These ratios were applied to the cases in a weighted least squares regression analysis to adjust for potential clinician bias. The relationships between dependent variables thought to be associated with clinical spinal stenosis (velocity in a 15-min walk test, pain measured on a 10-cm visual analog scale, and the pain disability index) and the independent variable of paraspinal mapping score were estimated.

RESULTS

A total of 150 subjects enrolled in the initial phase of the study. The mean age was 65.1 yrs (SD, 7.5; range, 55–79 yrs), with 38.7% male \((n = 58)\), 90.6% white, 5.3% black, and 4.0% Asian. The body mass index, calculated as weight in kilograms divided by squared height in meters, averaged 28.7 (SD, 5.9; range, 16.7–46.2). Eighty-eight individuals returned for the follow-up study 18 mos later. Three subjects were not tested for EMG (two in the initial phase and one at follow-up), resulting in 235 cases available for the current study.

Table 1 shows that when the examiner thought the subject had mechanical back pain (a disorder presumed not to involve the spinal nerves), they gave significantly lower scores than their masked colleagues. People thought to be asymptomatic tended to score less, and those thought to have stenosis tended to score higher.

The ratio of the examiner’s score to those of all other clinicians who tested the same subjects in
the same place shows substantial differences, as noted in Table 2. Examiners tended to score 0.17–3.7 times the group means, with some individuals showing statistically significant biases. The interclass correlation tells how well the examiner varied in the same way the group did. Notably, the tightest interclass correlation was for examiner A, the physician (A.H.) who had trained all of the other examiners in the technique.

Table 3 looks at paraspinal mapping along with the needle examination of the limb for fibrillation potentials and motor units. Significant physician bias was found in the analysis of motor units in all three populations. For abnormal spontaneous ac-

### Table 1

<table>
<thead>
<tr>
<th>Population</th>
<th>Examiners</th>
<th>Mean (SD)</th>
<th>Ratio</th>
<th>Paired Samples t Test</th>
<th>Intraclass Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unmasked Mean (SD)</td>
<td>Masked Mean (SD)</td>
<td>t</td>
<td>P</td>
<td>ICC 95% CI</td>
</tr>
<tr>
<td>Whole (n = 222)</td>
<td>0.55 (1.53)</td>
<td>0.75 (1.67)</td>
<td>-1.615</td>
<td>0.108</td>
<td>0.517</td>
</tr>
<tr>
<td>Physiatric initial diagnosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No pain (n = 61)</td>
<td>0.33 (1.11)</td>
<td>0.48 (1.35)</td>
<td>-0.652</td>
<td>0.517</td>
<td>-0.056</td>
</tr>
<tr>
<td>Back pain (n = 74)</td>
<td>0.28 (0.69)</td>
<td>0.86 (1.64)</td>
<td>-2.947</td>
<td>0.004</td>
<td>0.171</td>
</tr>
<tr>
<td>Stenosis (n = 87)</td>
<td>0.94 (2.12)</td>
<td>0.85 (1.89)</td>
<td>0.442</td>
<td>0.659</td>
<td>0.695</td>
</tr>
</tbody>
</table>

a ANOVA: paraspinal mapping score of unmasked examiners by physiatrist initial diagnosis: F = 4.808, P = 0.009.
b ANOVA: paraspinal mapping score of masked examiners by physiatrist initial diagnosis: F = 1.153, P = 0.318.
CI, confidence interval.

### Table 2

<table>
<thead>
<tr>
<th>Examiners</th>
<th>n</th>
<th>Mean (SD)</th>
<th>Ratio</th>
<th>Paired Samples t Test</th>
<th>Intraclass Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unmasked A; board certified</td>
<td>43</td>
<td>0.63 (1.98)</td>
<td>1.575</td>
<td>1.151</td>
<td>0.256</td>
</tr>
<tr>
<td>Masked all others</td>
<td>43</td>
<td>0.40 (1.38)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unmasked B; fellow</td>
<td>20</td>
<td>1.45 (2.33)</td>
<td>2.900</td>
<td>2.298</td>
<td>0.033</td>
</tr>
<tr>
<td>Masked all others</td>
<td>20</td>
<td>0.50 (1.10)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unmasked C; fellow</td>
<td>27</td>
<td>0.37 (1.39)</td>
<td>0.322</td>
<td>-2.294</td>
<td>0.030</td>
</tr>
<tr>
<td>Masked all others</td>
<td>27</td>
<td>1.15 (2.71)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unmasked D; board certified</td>
<td>2</td>
<td>1.00 (1.41)</td>
<td>2.00</td>
<td>0.333</td>
<td>0.795</td>
</tr>
<tr>
<td>Masked all others</td>
<td>2</td>
<td>0.50 (0.71)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unmasked E; board certified</td>
<td>43</td>
<td>0.26 (1.09)</td>
<td>0.295</td>
<td>-2.058</td>
<td>0.046</td>
</tr>
<tr>
<td>Masked all others</td>
<td>43</td>
<td>0.88 (1.72)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unmasked F; fellow</td>
<td>32</td>
<td>0.28 (0.85)</td>
<td>0.475</td>
<td>-1.380</td>
<td>0.177</td>
</tr>
<tr>
<td>Masked all others</td>
<td>32</td>
<td>0.59 (1.50)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unmasked G; fellow</td>
<td>22</td>
<td>0.86 (1.36)</td>
<td>1.049</td>
<td>0.097</td>
<td>0.923</td>
</tr>
<tr>
<td>Masked all others</td>
<td>22</td>
<td>0.82 (1.53)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unmasked H; fellow</td>
<td>22</td>
<td>0.23 (0.75)</td>
<td>0.174</td>
<td>-2.935</td>
<td>0.008</td>
</tr>
<tr>
<td>Masked all others</td>
<td>22</td>
<td>1.32 (1.64)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unmasked I; fellow</td>
<td>10</td>
<td>1.10 (2.13)</td>
<td>3.667</td>
<td>1.078</td>
<td>0.309</td>
</tr>
<tr>
<td>Masked all others</td>
<td>10</td>
<td>0.30 (0.48)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ANOVA for unmasked examiners: F = 1.932, P = 0.066.
ANOVA for masked examiners: F = 1.466, P = 0.171.
### Table 3A
Masked physician biases for paraspinal mapping, limb spontaneous activity, and limb motor unit morphology in asymptomatic subjects. Spontaneous activity graded 0–4, motor unit graded 0–10/10 polyphasic (“polys” in the table) motor units.

<table>
<thead>
<tr>
<th>Examiners</th>
<th>Paraspinal Mapping $^a$</th>
<th>Mean Spontaneous Activity $^b$</th>
<th>Maximum Spontaneous Activity $^c$</th>
<th>Mean Polys $^d$</th>
<th>Maximum Polys $^e$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n$</td>
<td>Mean (SD)</td>
<td>Min</td>
<td>Max</td>
<td>$n$</td>
</tr>
<tr>
<td>A</td>
<td>10</td>
<td>1.80 (3.36)</td>
<td>0</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>B</td>
<td>13</td>
<td>1.38 (1.90)</td>
<td>0</td>
<td>6</td>
<td>13</td>
</tr>
<tr>
<td>C</td>
<td>3</td>
<td>0.33 (0.58)</td>
<td>0</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>D</td>
<td>13</td>
<td>4.62 (6.46)</td>
<td>0</td>
<td>19</td>
<td>13</td>
</tr>
<tr>
<td>E</td>
<td>9</td>
<td>0.44 (0.73)</td>
<td>0</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>F</td>
<td>2</td>
<td>4.00 (5.66)</td>
<td>0</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>G</td>
<td>6</td>
<td>1.17 (1.33)</td>
<td>0</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>H</td>
<td>3</td>
<td>2.33 (4.04)</td>
<td>0</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>I</td>
<td>2</td>
<td>2.50 (0.71)</td>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>61</td>
<td>2.10 (3.79)</td>
<td>0</td>
<td>19</td>
<td>61</td>
</tr>
</tbody>
</table>

$^a$ $F = 1.227$, $P = 0.392$.  
$^b1$ $F = 0.794$, $P = 0.610$.  
$^b2$ $F = 1.289$, $P = 0.270$.  
$^c1$ $F = 3.669$, $P = 0.002$.  
$^c2$ $F = 4.554$, $P = 0.000$.  

### Table 3B
Masked physician biases for paraspinal mapping, limb spontaneous activity, and limb motor unit morphology in back pain subjects. Spontaneous activity graded 0–4, motor unit graded 0–10/10 polyphasic (“polys” in the table) motor units.

<table>
<thead>
<tr>
<th>Examiners</th>
<th>Paraspinal Mapping $^a$</th>
<th>Mean Spontaneous Activity $^b$</th>
<th>Maximum Spontaneous Activity $^c$</th>
<th>Mean Polys $^d$</th>
<th>Maximum Polys $^e$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n$</td>
<td>Mean (SD)</td>
<td>Min</td>
<td>Max</td>
<td>$n$</td>
</tr>
<tr>
<td>A</td>
<td>19</td>
<td>3.05 (5.21)</td>
<td>0</td>
<td>23</td>
<td>20</td>
</tr>
<tr>
<td>B</td>
<td>8</td>
<td>1.38 (2.20)</td>
<td>0</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>C</td>
<td>3</td>
<td>0.00 (0.00)</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>D</td>
<td>16</td>
<td>5.19 (7.59)</td>
<td>0</td>
<td>22</td>
<td>16</td>
</tr>
<tr>
<td>E</td>
<td>15</td>
<td>0.60 (1.35)</td>
<td>0</td>
<td>5</td>
<td>15</td>
</tr>
<tr>
<td>F</td>
<td>6</td>
<td>4.50 (8.71)</td>
<td>0</td>
<td>22</td>
<td>9</td>
</tr>
<tr>
<td>G</td>
<td>4</td>
<td>2.25 (3.30)</td>
<td>0</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>H</td>
<td>1</td>
<td>3.00 (---)</td>
<td>3</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>I</td>
<td>2</td>
<td>5.50 (0.71)</td>
<td>5</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>74</td>
<td>2.85 (5.32)</td>
<td>0</td>
<td>23</td>
<td>80</td>
</tr>
</tbody>
</table>

$^a$ $F = 1.058$, $P = 0.403$.  
$^b1$ $F = 3.009$, $P = 0.006$.  
$^b2$ $F = 3.678$, $P = 0.001$.  
$^c1$ $F = 3.939$, $P = 0.001$.  
$^c2$ $F = 4.845$, $P = 0.000$.  

---

tivity, there was little opportunity for variation in the asymptomatic subjects, who were almost all without fibrillations. In the back pain group, where some fibrillations were detected, interexaminer differences were statistically significant. Among the people with spinal stenosis, there was an almost significant trend for physician bias regarding the maximum scores. In contrast, for all three groups, no significant trend toward bias was noted for paraspinal mapping.

Of 11,280 paraspinal locations tested in this trial (48 per subject × 235 cases), a total of 439 (3.89%) produced invalid scores, resulting in 66 (3.51%) invalid spinal levels and 13 (5.53%) invalid cases. Table 4 shows the details of this analysis. These activities varied significantly among the examiners, all statistically P < 0.001. The post hoc analysis using Tukey's honest significant difference test identified one particular examiner (F) as having a significantly higher percentage of invalid scores and of invalid levels compared with examiners A, D, E, and G (P < 0.01) and a significantly higher percentage of invalid cases compared with examiners A, D, and E. The fellows had more invalid cases compared with the faculty (n = 102, 11.76 ± 32.38% invalid cases for the fellows vs. n = 133, 0.75 ± 8.67% for the faculty, t = 3.753, P < 0.001). The only fellow whose ability to obtain valid insertional activity was similar to the faculty physicians had a year of clinical experience before the fellowship.

Table 5 shows the ability of EDX scores to predict outcome variables—ambulation velocity, overall disability, and pain—before and after weighting. Paraspinal mapping seems to explain more variation in the 15-min walking speed test than in the other two tests. When statistical adjustments were made for physician bias (from data in Table 2) the relationships between paraspinal mapping and the pain/function measures improved.

**DISCUSSION**

The credibility of medical consultation resides in both the technical component of the examination and the clinician’s interpretation of findings. The tension between the objectivity of testing and the requirement for clinical judgment is not well understood, but it is critical to the care of patients whose treatment plan depends on a diagnostic consultation rather than a simple test. For simple tests, a number of well-spelled-out factors come in to play.14-20 Here, we also validate some of the most basic tenets of science as they apply to testing within a consultation: unmasked observations can be biased21; different observers see the same phenomenon differently;22,23 elimination of variation improves quality;24,25 and quality improves with skill.26-28 The final finding—that correcting for bias improves the relationship of test results with the actual function of the patient—
means that these observations have clinical relevance. Specific hypotheses are discussed below:

1. Subjectivity is related to clinical presentation. The data in Table 1 show that the unmasked clinician scored people he or she thought had stenosis higher than a masked colleague; people with no symptoms were scored lower, and people thought to have only mechanical back pain were scored lower than they were by the masked colleague. Research in any number of scientific areas points to the inability of unmasked evaluators to give unbiased results. The American Association for Electrodiagnostic and Neuromuscular Medicine task force on blinded research recently raised concern about the need for masking, something not done adequately in any trial before the current one. Although our data show that research on technical electrodiagnostic data should be masked, other research on the practice of electrodiagnostic medicine should not necessarily “dumb down” the consultation by masking and fixed protocols, simply because masking is an important technical factor.

2. Examiners differ significantly in scoring people with the same spinal disorder. Expertise in all fields involves a process of education and experience tempered by the personality of the practitioner. It is not surprising that different electrodiagnosticians are more or less conservative in their testing. Liberal labeling of fibrillations that are in fact not truly reproducible is a serious problem for neophytes. On the other hand, careful or conservative interpretation of findings is equally misleading in comparison with established norms.

It is noteworthy that examiner A, who instructed the other testers on the paraspinal mapping technique, had scores that best related to the other examiners. This finding suggests that better instruction could further eliminate bias, pointing toward the influence of instructors on the habits of others. A written or video competency examination or observation of more cases might result in better consistency.

3. Subjectivity can be minimized by more codified and quantitatively validated protocols. A basic tenet of business, if not medicine, is that elimination of variation results in an improvement in quality. Results show that the interrater variability of paraspinal mapping is less than for any other component of the EMG examination. The limb motor unit examination as performed here had substantial variation. Not unexpectedly or coincidentally, our research elsewhere shows that the clinical diagnosis of spinal stenosis relates most strongly to the paraspinal mapping score, less so to the limb fibrillations, and only weakly to the limb motor unit examination, both of which are less well codified.

4. Skill impacts the validity of results. Table 4 shows substantial variation in the clinicians’ abilities to obtain valid results. Most remarkably, when this information is divided by experienced vs. inexperienced clinicians, there is no overlap. These fellows were not neophytes. They had performed a

<table>
<thead>
<tr>
<th>Number of subjects</th>
<th>Invalid scores of 48 per subject</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>235</td>
</tr>
<tr>
<td>Invalid scores</td>
<td>49</td>
</tr>
<tr>
<td>Invalid %</td>
<td>13</td>
</tr>
<tr>
<td>Invalid levels</td>
<td>38</td>
</tr>
<tr>
<td>Invalid %</td>
<td>17</td>
</tr>
<tr>
<td>Invalid cases</td>
<td>9</td>
</tr>
<tr>
<td>Invalid %</td>
<td>9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Invalid levels (zeroes in all six cells; eight levels per subject)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
</tr>
<tr>
<td>Invalid levels</td>
</tr>
<tr>
<td>Invalid %</td>
</tr>
<tr>
<td>Invalid cases</td>
</tr>
<tr>
<td>Invalid %</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Invalid cases (two levels on a side absent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
</tr>
<tr>
<td>Invalid levels</td>
</tr>
<tr>
<td>Invalid %</td>
</tr>
<tr>
<td>Invalid cases</td>
</tr>
<tr>
<td>Invalid %</td>
</tr>
</tbody>
</table>

* Three cases that had no EMG testing were excluded from the analysis. Faculty: examiners A, D, and E; fellows: examiners B, C, F, G, H, and I.
few hundred EDX examinations before their fellowship, and they were involved in intensive advanced EDX training. Additional evidence on the effects of experience and training is the recent work by Dillingham and Pezzin30 showing that nonphysician electrodagnosticians frequent miss the opportunity to test for neuropathy in diabetics, whereas specialist physicians do not. The claims that non-specialists can adequately practice electrodagnostic medicine seem questionable in light of this continuing learning curve among experts.

5. Adjusting for subjectivity improves results. Subjectivity of testing results in a poorer relationship between EMG test results and clinical information. All of the work on hypotheses 1–4 means little if the findings do not affect the clinical diagnosis. Table 5 shows that at least one factor, examiner bias, had a negative impact on the relationship between the EDX testing and the clinical picture. Adjusting for this bias improved the relationships between EMG scores and functional status.

The Michigan Spinal Stenosis Study methodology has a number of strengths that support our conclusions. It is prospective; it includes masked and unmasked examiners, different clinical populations, different physician training and experience, and codified and uncodified test components; and it has a large number of subjects. The study did not test variability for the whole paraspinal mapping grid—only one level was retested. In EDX, the extent of testing relates to medical complexity and disease severity, not just the presenting complaint.31 In this protocol, however, other clinical decisions—for example, the number of muscles or nerves to be tested7—were fixed protocol. Variations in the extent of exploration of muscles, physician professionalism, and attention to detail, and the possibilities of neglect, incompetence, and fraud are important variables to be considered if referral sources are to trust a consultation.

Most improvement in reproducibility will not come as a result of mere vigilance on the part of the individual clinician. It is an issue of consensus between different clinicians. Thus, professional organizations should develop standards regarding the factors that can improve consistency. Some of these factors have been researched. Others must be standardized through consensus or even arbitrary definition until research justifies certain choices. But without standardization across practices, variation will decrease the value of consultation.

**CONCLUSIONS**

Clinicians and scientists in all areas of diagnostic consultation need to understand that examiner subjectivity and variation do occur and can affect seemingly technical components of the consultation. More codification of tests and more study
of subjective factors will improve the utility of diagnostic consultation.

ACKNOWLEDGMENTS

Study examiners and staff also included April M. Petzer, DO, Marcus J. Harris, BS, Janis Huff, Richard W. Kendall, DO, Christopher M. Parres, BS, Allan Rowley, MD, Matthew J. Smith, MD, Andre Taylor, MD, and John A. Yarjanian, DO.

REFERENCES


Fatigue-Induced Changes in Phasic Muscle Activation Patterns During Dynamic Trunk Extension Exercise

ABSTRACT

Objective: To investigate the influence of fatigue on phasic muscle-activation patterns during dynamic trunk extension exercise.

Design: Fifteen healthy volunteers performed dynamic trunk-extension exercise through a 30-degree range-of-motion (ROM) exercise to task failure at an intensity of 50% of maximum. Electromyography (EMG) signals were recorded unilaterally from the lumbar extensor, gluteus maximus, and biceps femoris muscles, and signal amplitude was analyzed in 10-degree increments during the unfatigued and fatigued states (0–10 degrees from torso horizontal to the ground was considered extension, and 11–20 and 21–30 degrees of flexion relative to this were considered midphase and flexion, respectively).

Results: Lumbar extensor EMG was approximately 75% of maximum EMG, with no differences being observed with respect to ROM or fatigue state. The gluteus maximus demonstrated an altered phasic activation pattern with fatigue, with an increased recruitment during the extension phase (fatigued-state extension-phase EMG: 89.1 ± 8.3% > flexion phase EMG: 37.8% ± 9.1%). The biceps femoris demonstrated a similar response during both the fatigued and unfatigued states (fatigued-state extension EMG: 77.8 ± 5.4% > midphase EMG: 65.8 ± 5.7% > flexion EMG: 46.8 ± 4.0%; unfatigued-state extension EMG: 46.1 ± 3.7% > flexion EMG: 27.1 ± 2.6%).

Conclusions: During this exercise, as one moves from flexion to extension, hip extensor muscle activity increases, whereas lumbar extensor activity does not. Additionally, fatigue results in an altered recruitment pattern, with the hip extensors being activated to a greater extent in the extension phase. These findings suggest that when this exercise is performed in the prone position, it can be used to stimulate the lumbar and hip extensor muscles, but the specific exercise protocol in terms of set/repetition number and ROM will influence which muscles are primarily targeted.

Key Words: Muscle Activation, EMG, Fatigue, Exercise, Lumbar, Trunk Extension
Weak and fatigable lumbar extensor muscles are reported in patients suffering from low-back pain. Therefore, the development of muscular strength and endurance in the trunk extensor muscles are believed to be important in the treatment and rehabilitation of low-back pain and injury. Consequently, the U.S. Agency for Health Care Policy and Research has concluded that back extensor muscle conditioning exercises are helpful in the management of acute low-back pain.

Although numerous devices and exercise protocols have been developed to strengthen the lumbar extensor muscles, there is not a standard exercise prescription to prevent or treat low-back pain. One exercise commonly performed to exercise the lumbar paraspinal muscles is trunk extension in the prone position (Fig. 1). To serve in improving this prescription, the complex activation patterns of the back extensor musculature during compound trunk-extension exercise needs to be better understood. Because the lumbar spine is tightly coupled to the gluteus maximus and biceps femoris muscles via the thoracolumbar fascia and ligamentum sacrotuberale, the hip extensor muscles can contribute to the force production if pelvic derotation (movement of the pelvis into anterior rotation relative to the femur) is permitted. Additionally, the induction of muscle fatigue has been shown to alter the synergistic relationship of these muscles, with the activity of the gluteus maximus increasing to a disproportionately large extent. At present, the muscle-activation patterns at different points throughout the range of motion (ROM) during fatiguing trunk-extension exercise performed in the prone position are not known. Therefore, the purposes of this study were to evaluate the phasic muscle-activation patterns of the lumbar and hip extensors at different points of movement while performing dynamic trunk-extension exercise, and to determine the influence of muscle fatigue on these activation characteristics. Because this exercise mode permits pelvic derotation, we hypothesized that the hip extensor muscle activity would increase as the trunk approached the horizontal position during the exercise. Furthermore, we hypothesized that muscle fatigue would alter these patterns, allowing the hip extensors to contribute to the force production to an even greater extent during a fatigued state.

METHODS

Subjects

Nine male and six female volunteers were recruited from a university setting to participate in this study. Descriptive statistics are provided in Table 1. Subjects were apparently healthy and recreationally active but were not currently engaged in a systematic exercise program for the lumbar or hip extensor muscles. The rights of human subjects were protected in that the experimental protocol was reviewed and approved by the university’s institutional review board, and before testing, subjects provided written informed consent. Potential subjects were excluded if they had a history of chronic low-back pain, present back pain, or orthopedic or cardiovascular contraindications to exercise.

Experimental Design

Data were collected in the musculoskeletal research laboratory at Syracuse University. Each subject reported to the lab for a familiarization session, at which time height, body mass, and upper-body mass were recorded. Additionally, the trunk-extension protocol was described, and subjects practiced the movement. On a subsequent visit, subjects performed the trunk-extension fatigue protocol. During this visit, lumbar strength was determined, and subjects performed isotonic trunk-extension exercise through a 30-degree ROM at 50% maximum voluntary contraction force (MVC) on a variable-angle Roman chair (BackStrong International, Berea, CA) positioned at 15 degrees relative to horizontal (Fig. 1). Zero-degree ROM was noted when the subject’s torso was horizontal to the ground, and the movement consisted of 30 degrees of flexion from this point. Subjects were asked to perform as many repetitions as possible, and the test was terminated when a subject could no longer complete the 30-degree ROM or complete the repetitions in the prescribed time (2 secs concentric, 2 secs eccentric). During exercise, muscle activity was recorded from surface electromyography (EMG) electrodes located on the right lumbar paraspinal (L4–L5), gluteus maximus, and biceps femoris muscles. An electric

![FIGURE 1](image-url)
Goniometer located on the right hip provided data for the trunk relative to the thigh position. Detailed information on all procedures is described below.

**Determination of MVC Force**

To determine lumbar strength, subjects were fitted with a nylon torso harness equipped with a ring at the midsternal region to allow for a chain attachment. Next, they were positioned on a variable-angle Roman chair at 15 degrees relative to horizontal, with the upper body horizontal to the ground (0 degrees) and then attached to a tensiometer (Takei Scientific Instruments, Tokyo, Japan) via the harness and chain. Subjects crossed their arms and placed their hands on the opposite shoulders. During the strength assessment, subjects gradually increased force production over the first second and then exerted a maximum effort for 2–3 secs. Three maximal contractions were performed with a 2- to 3-min rest period between efforts. If subjects continually recorded more force with increasing trials, or if the trials were not within 2 kg, additional trials were performed until a plateau was reached. During testing, strong verbal encouragement was provided by the investigators. The trial resulting in the highest force production was considered maximal strength. The lumbar-extension strength-testing protocol has been previously described and found reliable (Pearson $r = 0.99$). Each subject’s MVC force was then determined as follows:

MVC force = strength (kg)  
+ harness and chain weight (0.7 kg)  
+ upper-body mass (kg)

To load the subjects at a relative intensity of 50%, the MVC force was multiplied by 0.50, and upper-body mass was subtracted. The resultant value was then added to the subjects upper body via a weight vest that was adjustable in 1.14-kg increments (WeightVest.com, Sugar City, ID). Strength and loading characteristics are displayed in Table 1.

**EMG**

EMG is commonly used to assess muscle fiber action potential activity in skeletal muscle. Expression of the EMG signal in the time domain allows for evaluation of neuromuscular activation patterns, because a greater amplitude seems to be primarily attributable to an increase in the number of motor units recruited and increased motor unit discharge rate.

Before testing, two square (16 square cm) Ag/AgCl surface EMG electrodes (Nikomed 2002 electrode, Danlee Medical Products Inc, Syracuse, NY) were placed on each subject’s right paraspinus region, 1 cm above and below the L4–L5 interspinous space. Additionally, EMG electrodes were placed over the right gluteus maximus and right biceps femoris muscles. Electrode placement was chosen according to Cram and Kasman’s standardized electrode-placement atlas. The gluteus maximus electrodes were placed at the midpoint of a line running from the inferior lateral angle of the sacrum to the greater trochanter. Biceps femoris muscle electrodes were placed in the middle of the distance between the gluteal fold and popliteal joint. A bipolar electrode configuration was used with an interelectrode distance of 25 mm. A reference electrode was placed on a bony protuberance (i.e., iliac crest). Before electrode application, the skin was shaved, abraded, and then cleaned with alcohol to minimize skin impedance.

---

**Table 1** Descriptive statistics and performance characteristics of subjects (means ± standard deviation)

<table>
<thead>
<tr>
<th>Age, yrs</th>
<th>Height, cm</th>
<th>Weight, kg</th>
<th>UBM, kg</th>
<th>MVC Force, kg</th>
<th>Added Load, kg</th>
<th>Repetitions to Failure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male, $n = 9$</td>
<td>22.2 ± 2.2</td>
<td>174.6 ± 5.8</td>
<td>78.4 ± 11.4*</td>
<td>37.2 ± 4.8*</td>
<td>69.3 ± 10.9*</td>
<td>19.4 ± 6.9*</td>
</tr>
<tr>
<td>Female, $n = 6$</td>
<td>21.1 ± 3.8</td>
<td>164.2 ± 4.0</td>
<td>65.5 ± 6.3</td>
<td>29.8 ± 4.6</td>
<td>45.5 ± 5.5</td>
<td>7.9 ± 4.0</td>
</tr>
</tbody>
</table>

UBM, upper-body mass; MVC, maximum voluntary contraction. Added load equals weight added to the upper body, to load individuals at a relative intensity equal to 50% of maximum strength. *Males > females ($P < 0.05$).
The interference EMG signal was recorded at a sampling rate of 1000 Hz. The signal was preamplified 100 times with a BioAmp 100 (Axon Instruments Inc., Foster City, CA). After preamplification, the signal was amplified 10 times with the use of a Cyber Amp 380 (Axon Instruments Inc., Foster City, CA) (total gain of 1000) and digitized with an analog-to-digital board via a LabView Data Acquisition Card (National Instruments, Austin, TX). The signal was band pass filtered between 10 and 600 Hz, and EMG signals were stored on disk for subsequent analysis by the Acqknowledge software program (Biopac Systems Inc., Santa Barbara, CA). For the first and last four contractions, the root mean squared (RMS) EMG was determined in 10-degree increments for the concentric portion of the movement according to the goniometer data. EMG values were then normalized to the maximal muscle activity recorded during the MVC testing (RMS EMG determined in a 1-sec period during the middle of the contraction).

To assure that maximal activity was obtained for the hip extensors, subjects were also asked to maximally contract the gluteus maximus and biceps femoris muscles for 4 secs. For the gluteal musculature, subjects stood with their feet shoulder-width apart and contracted the gluteal muscles. For the biceps femoris muscles, subjects were placed in a knee-extension/flexion dynamometer (MedX, Okala, FL) with the knee positioned at 90 degrees. For both muscles, practice trials were given along with strong verbal encouragement, and the task was performed two times. RMS EMG was determined for a 1-sec portion of each contraction. Three of the fifteen subjects obtained a greater muscle activation for the biceps femoris using this technique (when compared with the Roman chair MVC testing). Only one subject had increased activity for the gluteus maximus using this technique. The maximal muscle activity obtained from any of the three testing procedures were used for normalization.

**Fatigue Protocol**

Subjects performed isotonic trunk extensions to task failure. They were instructed to begin with the trunk in the flexed position and to extend their trunks in a smooth, controlled fashion, completing the concentric phase of the dynamic exercise in a 2-sec period. Next, the subjects were instructed to lower their torsos during the eccentric phase in 2 secs to return to the flexed position. A metronome, along with investigator feedback, ensured appropriate timing. The exercise was performed through a 30-degree ROM, with extension being parallel to the ground. During exercise, subjects were verbally encouraged to perform as many repetitions as possible to task failure. Task failure was defined as the point at which the subject could no longer complete the ROM or follow the appropriate timing.

**Treatment of the Data**

For each muscle, regardless of the number of repetitions performed, RMS EMG was determined for the first and last four repetitions in 10-degree increments for the concentric portion of the exercise (0–10 degrees from the horizontal position = extension phase; 11–20 degrees = midphase; 21–30 degrees = flexion phase). The RMS EMG values for each phase were averaged for the first and last four repetitions and were normalized to the maximal RMS EMG recorded during MVC testing. These values were used to provide an indication of muscle activation at the various ROM phases during the unfatigued (first four repetitions) and fatigued (last four repetitions) muscle states.

Normalized EMG activity for each muscle was analyzed using a two-way factorial analysis of variance for repeated measures. Main effects for this statistical model were ROM phase (extension, mid-phase, and flexion) and muscle fatigue state (unfatigued and fatigued). Statistical significance was accepted at $P < 0.05$. All data are reported as means ± standard errors unless otherwise noted. The STATA statistical package (StatCorp Corporation, College Station, TX) was used for all data analysis.

**RESULTS**

**Lumbar Extensors**

No differences in lumbar paraspinal EMG were observed between the ROM phases or fatigue state (Fig. 2).

**Gluteus Maximus**

For the gluteal musculature, am ROM–fatigue state interaction was detected. Further analysis revealed the fatigued state resulted in significantly

![FIGURE 2 Lumbar extensor muscle activity during Roman chair trunk-extension exercise for the unfatigued and fatigued state of the muscle with respect to range of motion.](image-url)
Increased EMG levels at each phase of the ROM when contrasted to the unfatigued state (Fig. 3). Additionally, in the fatigued state EMG was higher during the extension phase vs. the flexion phase (extension phase EMG: 89.1 ± 8.3%; flexion phase EMG: 37.8 ± 9.1%) (Fig. 3). No differences were observed in phasic activation patterns in the unfatigued state (Fig. 3).

**Biceps Femoris**

For the biceps femoris an ROM–fatigue state interaction was also found. Further analysis revealed that it followed a similar pattern as the gluteus maximus. During the fatigued state, EMG was significantly increased at each ROM phase compared with the unfatigued state (Fig. 4). Muscle fatigue increased the activity of the biceps femoris from flexion to extension (midphase EMG > flexion EMG; extension EMG > flexion EMG) (Fig. 4). Interestingly, the biceps femoris was the only muscle demonstrating a significantly higher EMG activity with respect to ROM phase during the unfatigued state (extension EMG: 46.1 ± 3.7% > flexion EMG: 27.1 ± 2.6% (Fig. 4).

**DISCUSSION**

The most novel findings of this study are that (1) as one moves from flexion to extension, the hip extensors’ muscle activity increases while the lumbar paraspinal activity does not; and (2) fatigue induces alterations in phasic muscle-activation patterns during trunk-extension exercise, with an even greater reliance on the hip extensors at all joint angles, especially the extension phase of movement.

Few studies have evaluated the muscle activity of the hip extensors during trunk-extension exercise in the prone position. This study is in agreement with these studies in illustrating the importance of the hip extensors to trunk-extension performance. Our finding of gluteal muscle activity in the range of 25% (unfatigued state) of maximum is similar to the values reported by Arokoski and colleagues. More interesting is our observation of fatigue resulting in alterations to the phasic patterns, with a disproportionate increase in the gluteus maximus activity occurring in the extension phase, with mean muscle activity approaching nearly 90% of maximum. This finding suggests that trunk extension in the prone position can provide a significant stimulus to the gluteal musculature. Additionally, during the fatigued state, the biceps femoris demonstrates a progressive increase in muscle activity as one moves from flexion to extension. Therefore, these data suggest that during dynamic trunk-extension exercise in the prone position, as one becomes fatigued and moves toward extension, the hip extensors are activated to a greater extent to accommodate the workload, thus allowing the subjects to complete the full ROM.

Another interesting finding of this study is that the lumbar extensor muscle activity is unaltered with this fatiguing task. The observation that the lumbar musculature does not demonstrate an increase in the EMG signal associated with repetitive trunk-extension exercise is in disagreement with the typical response of most skeletal muscles. Other studies have indicated that the lumbar musculature response to fatigue and loading is unique in this aspect. Our finding that lumbar extensor muscle activity does not change from before to after task failure, with a concomitant in-
crease in hip extensor activity, is in agreement with previous findings from our laboratory. 13, 14 Recently, many of the subjects from the present study were also participants in a study evaluating a de-recruitment threshold of the paraspinal muscles associated with the same fatigue task. 14 That study demonstrated that during trunk-extension exercise in the prone position (identical to that performed in the present study), the activity of the lumbar musculature does increase significantly up to approximately 55% of the maximal number of repetitions one can perform, at which point decrements in lumbar muscle activity begin to occur. 14 Therefore, it is plausible that an increase in lumbar muscle activity did occur but that, with the onset of fatigue, the synergistic hip extensors were recruited and the lumbar muscles were derecruited.

The present study also indicates that lumbar muscle activity is unaltered with respect to the exercise ROM. Regardless of fatigue or ROM, the lumbar musculature was activated between approximately 70 and 75% of maximum. This finding is contrary to what is found in most skeletal muscles, where the magnitude of the surface EMG signal increases as a shortening contraction progresses. 15 This further illustrates the unique response of the lumbar musculature to this particular exercise.

The clinical implications of these findings suggest that if the goal of an exercise prescription is to train for strength and endurance of the lumbar extensors, the ROM will not alter the recruitment. However, because the gluteal and biceps femoris musculature are largely activated during trunk-extension exercise in the prone position, it is suggested that the movement be performed through the full ROM to aid in activating all of the muscles completely. Additionally, the finding of alterations with fatigue has clinical implications with respect to the number of sets and repetitions prescribed to a patient. For example, when these data are taken in consideration with our previous reports, 13, 14 it seems that if this trunk-extension exercise is being conducted primarily to benefit the lumbar extensors, then exercise to muscular failure does not result in an increased recruitment of the paraspinal muscles, at least not when the relative intensity is approximately 50% of peak force. However, if this trunk-extension exercise in the prone position is being used to stimulate the hip extensor musculature, a certain degree of fatigue (around 50% of maximum endurance) must be induced for these muscles to be further activated to contribute to a greater extent in the force production. It must be noted that these results and the resultant recommendations are from a cohort of young, healthy subjects, and may not be generalizable to low-back pain populations.

In summary, this study indicates that as one moves from a relative position of trunk flexion to extension, the gluteus maximus and biceps femoris muscle EMG activity increases, whereas the lumbar extensor activity does not. Additionally, muscle fatigue during this exercise results in an altered recruitment pattern, with the gluteal and biceps femoris muscles being activated to a greater extent in the extension phase. These findings suggest that this trunk-extension exercise in the prone position can be used to stimulate the lumbar extensors, gluteus maximus, and biceps femoris muscles, but the specific exercise protocol chosen in terms of fatigue state (i.e., number of sets and repetitions) and ROM will influence which muscles are primarily targeted.

REFERENCES


CME Self-Assessment Exam

Answers
American Journal of Physical Medicine & Rehabilitation
Vol. 86, No. 5 • May 2007

CME Article Number 1:
J. Bach, et al.

1. B
2. A
3. A
4. D
5. C
A Clinical Tool for Office Assessment of Lumbar Spine Stabilization Endurance: Prone and Supine Bridge Maneuvers

ABSTRACT


Objective: To assess the validity and reliability of an office-based surrogate measure of lumbar spine—stabilization endurance capability; to establish norms and reliability in an asymptomatic group; and to compare their measures with those from a group of chronic mechanical low-back pain patients.

Design: Eight healthy subjects participated in the tool-validation portion of the study that consisted of surface electromyographic (EMG) measurements of core muscle activation during prone and supine bridging. Subsequently, normative and test–retest reliability measures of prone and supine bridging duration were recorded from 43 subjects without back pain and were compared with those of 32 subjects with chronic mechanical low-back pain.

Results: Surface EMG indicated significantly preferential activation of anterior core muscles during prone bridging and posterior core muscles during supine bridging. Mean bridge durations for subjects without back pain were 72.5 ± 32.6 (mean ± SD) secs in prone and 170.4 ± 42.5 secs in supine. They were significantly less in subjects with back pain: 28.3 ± 26.8 secs in prone and 76.7 ± 48.9 secs in supine. Test–retest reliability using Pearson's correlation for prone and supine bridging was 0.78 and 0.84, respectively.

Conclusions: Bridging maneuvers seem to be practical, reliable, and valid methods of reflecting lumbar spine—stabilization endurance capability. Prone bridging preferentially challenges core flexors, whereas supine bridging recruits primarily the core extensors; both are compromised in patients with low-back pain.

Key Words: Core Stabilization, Bridging, Norms, Reliability, Validity
The management of chronic low-back pain is challenging, both from a medical and a societal perspective. Typically, a physician's office assessment of a patient with back pain complaints focuses on establishing a diagnostic category from which a treatment plan can be developed. In contrast, it is less common for physicians to perform back fitness testing in an office setting because of a number of constraints: lack of time, equipment, space, qualified personnel, and a paucity of validated techniques and norms.

One aspect of back fitness considered to be important in minimizing the severity and frequency of mechanical back pain is core stabilization.1–6 Core-stabilization exercise programs have been shown to be particularly effective in low-back pain patients with segmental hypermobility.7–9 Proper neuromuscular control, strength, and endurance of core muscles are thought to protect the spine by maintaining intervertebral neutral zones (defined as the part of the range of physiologic intervertebral motion, measured from the neutral position, within which the spinal motion is produced with a minimal internal resistance), within physiologic limits.10,11 The intervertebral neutral zone increases in size with injury and degenerative disease.12,13 It is likely that optimal core stability requires a combination of strategic motor control (particularly of the deep muscles such as multifidus and transversus abdominis) and adequate endurance capacity of the superficial muscles (rectus abdominis, obliques, erector spinae, quadratus lumborum).14–17 Therefore, a surrogate measure capable of reflecting core-stabilization capability may be a useful tool for physicians to identify patients requiring further rehabilitation and to monitor the progress of training efforts. To be practical and useful, the tool would need to be simple, valid, reliable, and able to be performed efficiently in an office setting without significant risk to the patient.

We identified two maneuvers described for core-stabilization training as being potentially well suited for this purpose because they seemed to be relatively simple to administer and perform: (1) the supine bridge, for assessment of trunk stability against a flexion moment; and (2) the prone bridge, for assessment of trunk stability against an extension moment. Advocates of these exercises claim that supine bridging selectively recruits the spinal extensors, whereas prone bridging recruits the flexors.18 However, these claims have not been validated, nor have the tests' reliability and normative values been established.

The purposes of this study were to assess the validity, reliability, and practicality of prone and supine bridging as office-based assessment tools designed to measure lumbar spine–stabilization endurance capability; to establish norms in an asymptomatic population; and to compare their values with those from a group of chronic mechanical low-back pain patients.

METHODS

Evaluations of the prone and supine bridging maneuvers were conducted in three phases: validation; reliability and norms from an asymptomatic group; and values from a symptomatic group. The study was approved by the University of Alberta ethics review board, and all subjects gave their informed consent.

Prone and Supine Bridging Maneuvers

Prone Bridge

Each subject began in the prone position, propped on the elbows. The elbows were spaced shoulder-width apart, and the feet were set with a narrow base, but not touching. The subject then raised the pelvis from the floor so that only the forearms and the toes were in contact with the floor. The shoulders, hips, and ankles were maintained in a straight line (Fig. 1A). The position was held until fatigue or pain prevented maintenance of the test position.

Supine Bridge

Each subject began in the supine position with knees flexed 90 degrees and the soles of the feet on...
the floor with a narrow base, but not touching. The thighs could not be in contact. The hands were positioned by the ears. The subject then raised the pelvis from the floor so that the shoulders, hips, and knees were maintained in a straight line (Fig. 1B). The position was held until fatigue or pain prevented maintenance of the test position. If the subject reached 2 mins, the dominant leg was extended at the knee, removing one point of support (Fig. 1C). This was designed to shorten the bridge duration by increasing the difficulty through addition of a torque moment to the core and an increase in the counterbalance weight. In the event of a preexisting injury to the dominant support limb (e.g., deficient anterior cruciate ligament), the nondominant leg was extended instead.

Validation

Eight subjects (six males, two females) were recruited by flyers posted in a rehabilitation hospital and underwent surface electromyographic (sEMG) assessment while performing the prone and supine bridging maneuvers. The target population was between the ages of 18 and 65 yrs, healthy, and without a recent history of back pain. Exclusion criteria included a history of angina, emphysema, shoulder pain in the past 6 mos, diagnosed spinal abnormality, abdominal or back surgery within the past year, cervical strain, and other pain conditions. The subjects were asked to complete a demographics form for the purpose of establishing age, weight, height, gender, and frequency of exercise in an average week.

After becoming familiarized with the test exercises, surface electrodes were placed over four muscle groups in a bipolar configuration on the right side of the body: rectus abdominis, external oblique, erector spinae, and hamstring. The rectus abdominis site was centered on the muscle belly midway between the pubis and the umbilicus. The external oblique site was 5 cm above the anterior superior iliac spine. The erector spinae site was 2 cm lateral to the L4–L5 interspace, and the hamstring site was at the midpoint of the muscle bellies. The hamstring was used instead of the gluteus maximus because of easier access and the technical advantage of less overlying adipose tissue. The skin was prepped by shaving the hair, applying gentle abrasion with an emery board, and cleansing with an alcohol swab. The paired electrodes were placed perpendicular to the muscle fibers, 3 inches apart. Sensitivity of the amplifier was set at 0.1 mV per division. Custom-written Labview software was used for data acquisition and analysis.

Maximum voluntary contractions were recorded for each muscle group by having each subject perform a maximal isometric contraction against resistance as follows:

1. Rectus abdominis: The subject was supine, with knees flexed at 90 degrees. The examiner restrained the shoulders and knees while the subject attempted to do a sit-up.
2. External obliques: The subject was supine, with knees flexed at 90 degrees. The examiner restrained the right shoulder and left knee while the subject attempted to bring the right shoulder up to the left knee.
3. Erector spinae: The subject was prone. The examiner restrained the shoulders and hips while the subject tried to raise the trunk up off the table.
4. Hamstrings: The subject was prone. The examiner held down the right leg while the subject tried to raise it off the table by extending from the hip.

After several minutes of rest, subjects performed the prone bridge maneuver and then the supine bridge maneuver, as described above. Data from the four muscle groups were recorded simultaneously. The sEMG waveforms were rectified, and the area under the portion of the sEMG envelope at the plateau region was measured. Because the time this could be sustained was different in each subject, the area was normalized to the duration of the segment. To make the sEMG output comparable between subjects, it was also expressed as a percentage of the electromyographic (EMG) output during the maximum voluntary contraction. The mean and standard deviation values are reported for each muscle group. A paired t test was used to compare the difference in the pattern of muscle activation with the two maneuvers for each muscle group.

Asymptomatic Group Norms and Reliability

Participants for this phase of the study were gathered from Hepburn, Saskatchewan, Canada. Flyers were posted in local businesses and bulletin boards. Forty-three subjects (22 males, 21 females) were recruited. Informed consent was obtained. Inclusion criteria were volunteers between the ages of 18–65 and no recent history of back pain. Exclusion criteria were as described in the validation portion of the study. After instruction and familiarization, each subject performed one repetition of the prone and supine bridging maneuvers on a portable gymnasium mat, in random order. Time to fatigue was measured in seconds. Reason for test cessation was recorded (pain vs. fatigue). A work to rest ratio of 1:4 was

382 Schellenberg et al.
observed. The maneuvers were then repeated to establish test–retest reliability. Statistical analysis was conducted using a paired t test to determine the average bridging times and to determine whether a test order effect was present. Pearson’s correlation coefficient was used to evaluate test–retest reliability of the two bridging maneuvers. Relationships between gender, reason for test cessation, and bridge duration were analyzed using one-way ANOVA.

Prone and Supine Bridge Values for Subjects with Low-Back Pain

Thirty-two subjects experiencing chronic mechanical low-back pain were recruited from the senior author’s practice (14 males, 18 females). Criteria for this diagnosis included low-back pain for >6 mos; no clinical or investigative evidence of spine infection, neoplasia, fracture, inflammatory disease, or neurologic impairment; pain aggravation with activity; and improvement with rest. After instruction and familiarization regarding the bridging exercises, each subject performed one repetition of the prone and supine bridge maneuvers. The test order of the maneuvers was randomized, and maneuver duration was recorded, as was each subject’s reason for cessation of the maneuver (primarily pain or fatigue). Descriptive statistics and Pearson’s correlation coefficient were used to summarize data and to explore relationships between bridging endurance, body mass index, age, pain (visual analog scale), and disability (Oswestry Disability Questionnaire Score). An unpaired t test was used to compare the bridging duration times of the asymptomatic and symptomatic groups. Body mass index was similarly compared. Demographic descriptions of the subjects from all three groups are shown in Table 1. Relationships between gender, reason for test cessation, and bridge duration were analyzed using one-way ANOVA.

RESULTS

Subject demographic characteristics of all three groups are summarized in Table 1.

Validation

From visual inspection of individual records, it was apparent that the prone bridge preferentially recruited the anterior core stabilizers, whereas the supine bridge preferentially recruited the posterior stabilizers (Fig. 2). Statistical assessment of the pooled data from all subjects on the EMG output

<table>
<thead>
<tr>
<th>TABLE 1 Subject demographics: mean (range)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pilot Data</td>
</tr>
<tr>
<td>n = 8</td>
</tr>
<tr>
<td>Age, yrs</td>
</tr>
<tr>
<td>Weight, kg</td>
</tr>
<tr>
<td>Height, cm</td>
</tr>
<tr>
<td>Exercise, hrs/wk</td>
</tr>
</tbody>
</table>

FIGURE 2 Surface electromyographic recordings from the flexors (rectus abdominis and external oblique) and the extensors (erector spinae and hamstrings) in a representative subject. Whereas the flexor muscles were vigorously activated during the prone bridge maneuver, the extensor muscles were quiescent. The reverse occurred during the supine bridge maneuver.
During maximum voluntary contraction, prone bridge, and supine bridge maneuvers confirmed this selective recruitment pattern (Table 2).

**Asymptomatic Group Norms and Reliability**

Prone and supine bridge durations (first and second repetitions and mean) are summarized in Table 3. Prone bridge durations were approximately 43% those of the supine bridge. Additionally, the duration of the second repetition of prone bridge was significantly less than the first, suggesting a test order effect. This was not seen for the supine bridge. The Pearson correlation coefficients for test–retest reliability of the prone and supine bridging tests were 0.78 and 0.84, respectively (*P* < 0.05). No significant correlation was found between bridge duration and subject age.

Males could maintain both prone and supine positions significantly longer than females. Additionally, the prone:supine ratio was significantly higher for the male subjects. Bridge duration times did not seem to be associated with a subject’s reason for stopping the maneuver (Table 4).

**Prone and Supine Bridge Scores for Subjects with Low-Back Pain**

Mean prone bridge duration for subjects with chronic mechanical back pain was 28.3 ± 26.8 secs, whereas supine bridge mean was 76.7 ± 48.9 secs. These were significantly shorter than for the asymptomatic control group (*P* < 0.001). Prone:supine ratio was 0.45 ± 0.43, which did not differ significantly from the asymptomatic control group. Bridge duration times were not significantly affected by gender or by a subject’s reason for stopping the maneuver. As with the asymptomatic group, prone:supine ratios were significantly higher in symptomatic male subjects (Table 5).

Prone bridging was significantly correlated with visual analog scale (*r* = −0.63) and Oswestry Disability Questionnaire Scores (*r* = −0.56). Supine bridging was significantly correlated with visual analog scale (*r* = −0.62) and Oswestry Disability Questionnaire Scores (*r* = −0.36). In contrast, no significant correlation was found between bridging endurance, age, and reasons for cessation. Body mass index was significantly higher in the symptomatic group (mean [SD] = 27.4 [3.7]) than the asymptomatic group (25.1 [4.2]; *t* = −2.6, *P* = 0.03). However, there was no significant correlation between body mass index and supine or prone bridging durations in either of the groups (supine: asymptomatic group *r* = −0.29, symptomatic group *r* = −0.35; prone: asymptomatic *r* = −0.24, symptomatic *r* = −0.24).

**DISCUSSION**

The first goal of this study was to assess the validity of the prone and supine bridge maneuvers as surrogate measures of lumbar spine–stabilization endurance. The sEMG assessment validated the claims that the prone bridge selectively recruits anterior trunk muscles and the supine bridge selectively recruits posterior trunk muscles. The muscle groups in this study were selected in an

<table>
<thead>
<tr>
<th>Muscle</th>
<th>Prone Bridge</th>
<th>Supine Bridge</th>
<th>t Statistic</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>%MVC Mean (SD)</td>
<td>%MVC Mean (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rectus abdominus</td>
<td>52.2 (22.5)</td>
<td>7.1 (4.4)</td>
<td>5.8</td>
<td>0.0006</td>
</tr>
<tr>
<td>External oblique</td>
<td>59.0 (26.6)</td>
<td>9.3 (4.0)</td>
<td>5.1</td>
<td>0.0013</td>
</tr>
<tr>
<td>Lumbar extensor</td>
<td>10.6 (3.4)</td>
<td>56.8 (19.7)</td>
<td>−7.5</td>
<td>0.0014</td>
</tr>
<tr>
<td>Hamstring</td>
<td>4.3 (2.4)</td>
<td>36.9 (16.2)</td>
<td>−5.73</td>
<td>0.0071</td>
</tr>
</tbody>
</table>

MVC, maximum voluntary contraction.

**TABLE 3 Prone and supine bridge durations in asymptomatic subjects**

<table>
<thead>
<tr>
<th></th>
<th>Mean, secs</th>
<th>Standard Deviation</th>
<th>t Statistic</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prone bridge 1</td>
<td>80.6</td>
<td>41.3</td>
<td>4.12</td>
<td>0.000175</td>
</tr>
<tr>
<td>Prone bridge 2</td>
<td>64.2</td>
<td>28.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>72.5</td>
<td>32.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supine 1</td>
<td>172.9</td>
<td>49.2</td>
<td>1.34</td>
<td>0.1881</td>
</tr>
<tr>
<td>Supine 2</td>
<td>167.5</td>
<td>39.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average of supine 1 and 2</td>
<td>170.4</td>
<td>42.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prone:supine ratio</td>
<td>0.43</td>
<td>0.17</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
attempt to evaluate the use and coordination of the core stabilizers. Endurance of these muscles has been shown to be more important than strength in the reduction of back pain and the prevention of future injury. In fact, only low levels of maximal voluntary contraction are required to ensure the stability of the spine in vivo. We used the hamstring muscles as a measure of hip extensor activity to show that the lumbar extensors are the primary contributors to stability in the supine bridge. The external oblique site was chosen to give an indication of anterolateral muscle activity in general, and the rectus abdominus site provided information about that muscle in particular.

Recently, strong emphasis has been placed on the importance of the timing and extent of activation of the deep local muscle system (particularly multifidi and transversus abdominus) by several investigators. The activity of these muscles during prone and supine bridge maneuvers is unknown but deserves further research. The ability of the prone and supine bridge maneuvers to discriminate between asymptomatic vs. symptomatic subjects suggests a level of concurrent validity. Additional concurrent validity is also suggested by the negative correlation between bridging duration and self-reported pain and disability scores among the subjects with mechanical low-back pain.

In this study, prone and supine bridging resulted in relatively high-percentage maximum voluntary contraction values, suggesting that the maneuvers may be useful not only for testing but also for training purposes. The high level of activation of the external oblique, a lateral stabilizer, during prone bridging is of clinical interest because it can potentially be trained using that maneuver. In addition to testing lateral stabilization endurance capability, McGill et al. also used side bridging to train the lateral stabilizer muscles.

Prone and supine bridging test–retest reliability in this study was good, but it was not as strong as that described by Hicks et al. The reliability of the prone and supine bridging test–retest reliability in this study was good, but it was not as strong as that described by Hicks et al. The reliability of the prone and supine bridging test–retest reliability in this study was good, but it was not as strong as that described by Hicks et al. The reliability of the prone and supine bridging test–retest reliability in this study was good, but it was not as strong as that described by Hicks et al. The reliability of the prone and supine bridging test–retest reliability in this study was good, but it was not as strong as that described by Hicks et al. The reliability of the prone and supine bridging test–retest reliability in this study was good, but it was not as strong as that described by Hicks et al. The reliability of the prone and supine bridging test–retest reliability in this study was good, but it was not as strong as that described by Hicks et al. The reliability of the prone and supine bridging test–retest reliability in this study was good, but it was not as strong as that described by Hicks et al. The reliability of the prone and supine bridging test–retest reliability in this study was good, but it was not as strong as that described by Hicks et al. The reliability of the prone and supine bridging test–retest reliability in this study was good, but it was not as strong as that described by Hicks et al. The reliability of the prone and supine bridging test–retest reliability in this study was good, but it was not as strong as that described by Hicks et al. The reliability of the prone and supine bridging test–retest reliability in this study was good, but it was not as strong as that described by Hicks et al. The reliability of the prone and supine bridging test–retest reliability in this study was good, but it was not as strong as that described by Hicks et al. The reliability of the prone and supine bridging test–retest relia…

### Table 4

<table>
<thead>
<tr>
<th></th>
<th>Mean, secs</th>
<th>Standard Deviation</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prone</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>51.2</td>
<td>19.9</td>
<td>29.5</td>
<td>0.0003</td>
</tr>
<tr>
<td>Male</td>
<td>92.9</td>
<td>29.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fatigue</td>
<td>72.3</td>
<td>33.0</td>
<td>0.36</td>
<td>0.55</td>
</tr>
<tr>
<td>Pain</td>
<td>79.1</td>
<td>33.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supine</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>152.0</td>
<td>30.2</td>
<td>0.36</td>
<td>0.55</td>
</tr>
<tr>
<td>Male</td>
<td>188.0</td>
<td>45.7</td>
<td>0.36</td>
<td>0.55</td>
</tr>
<tr>
<td>Fatigue</td>
<td>179.3</td>
<td>50.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain</td>
<td>155.6</td>
<td>20.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prone:supine ratio</td>
<td></td>
<td></td>
<td>13.9</td>
<td>0.006</td>
</tr>
<tr>
<td>Female</td>
<td>0.38</td>
<td>0.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>0.51</td>
<td>0.17</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 5

<table>
<thead>
<tr>
<th></th>
<th>Mean, secs</th>
<th>Standard Deviation</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prone</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>24.3</td>
<td>27.5</td>
<td>0.90</td>
<td>0.35</td>
</tr>
<tr>
<td>Male</td>
<td>33.4</td>
<td>26.0</td>
<td>1.4</td>
<td>0.25</td>
</tr>
<tr>
<td>Fatigue</td>
<td>30.6</td>
<td>26.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain</td>
<td>15.4</td>
<td>10.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supine</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>75.7</td>
<td>44.8</td>
<td>0.02</td>
<td>0.90</td>
</tr>
<tr>
<td>Male</td>
<td>77.9</td>
<td>55.4</td>
<td>1.4</td>
<td>0.25</td>
</tr>
<tr>
<td>Fatigue</td>
<td>73.5</td>
<td>47.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain</td>
<td>80.3</td>
<td>52.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prone:supine ratio</td>
<td></td>
<td></td>
<td>4.36</td>
<td>0.045</td>
</tr>
<tr>
<td>Female</td>
<td>0.31</td>
<td>0.22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>0.62</td>
<td>0.57</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
bridge maneuver may have been partly compromised by the test order effect, suggesting an element of fatigue with the second repetition because the amount of recovery time was likely inadequate.

The prone and supine bridging endurance times were significantly shorter for subjects with mechanical low-back pain. Although the symptomatic group was older than the asymptomatic normative group, age difference is unlikely to be an important influence on endurance time, as reflected by the lack of correlation between bridge duration and subject age in both the asymptomatic and symptomatic groups.

Overall, the prone and supine bridge maneuvers were well tolerated by the subjects with mechanical low-back pain. Although some had to terminate the maneuvers because of pain, it is interesting that the bridge duration was comparable regardless of whether the primary reason of cessation was pain or fatigue. Although the bridging endurance times were significantly shorter among symptomatic subjects, prone:supine ratios were not significantly different from the asymptomatic group, suggesting that anterior to posterior muscle imbalance may have a minimal relationship to mechanical low-back pain.

In terms of practicality, we aimed to describe a simple, quantifiable bedside measure of trunk-stabilizer endurance. Such a test should require a minimum of equipment, time, staff, and patient training. Ideally, the clinical tool can be used for both screening and monitoring purposes. The prone bridge and the supine bridge described in this study seem to fit those criteria. All testing for this study was done on a standard clinic examination table or a gymnasium mat and could be completed in approximately 5 mins.

REFERENCES
Factors Associated with Health-Related Quality of Life in Chronic Spinal Cord Injury

ABSTRACT


Objective: An important goal of rehabilitation and treatment after spinal cord injury (SCI) is to improve function and enhance health-related quality of life (HRQoL). However, previous assessments are limited by use of HRQoL instruments not specific to SCI. Although respiratory dysfunction is common in SCI, it has not been possible to assess the association of comorbid medical conditions, including respiratory symptoms and pulmonary function, to HRQoL. Therefore, we assessed whether these factors were associated with HRQoL in SCI using an SCI-specific HRQoL questionnaire.

Design: In our cross-sectional study, 356 participants ≥1 yr post-SCI completed a 23-item SCI-specific HRQoL questionnaire and a detailed health questionnaire, and underwent pulmonary function testing and a neurological exam at VA Boston between 1998 and June 2003.

Results: In a multivariate regression model, age, employment status, motor level and completeness of injury, and ambulatory mode (use of hand-propelled or motorized wheelchair, use of crutches or canes, or walking independently) were independently associated with HRQoL. After adjusting for these factors, chronic cough, chronic phlegm, persistent wheeze, dyspnea with activities of daily living, and lower forced expiratory volume in 1 sec and forced vital capacity were each associated with a lower HRQoL.

Conclusions: These results provide evidence for the clinical validity of our SCI-specific HRQoL instrument. We also identify potentially modifiable factors that, if addressed, may lead to HRQoL improvement in SCI.

Key Words: Quality Of Life, Spinal Cord Injuries, Pulmonary Function Tests, Comorbidity
Survival of individuals with chronic spinal cord injury (SCI) has improved, and coping with long-term permanent disability along with aging is an important part of day-to-day living for many injured persons. As a result, the goals of rehabilitation and treatment in SCI are to improve function and enhance health-related quality of life (HRQoL). Although specific definitions vary, HRQoL is a patient-centered outcome that relates to the physical, social, and psychological aspects of an individual’s well-being that may be affected by the burden of disease, disability, and related treatment. When assessed accurately, HRQoL measures are reliable, valid, and responsive to important clinical changes. It is essential to identify determinants of HRQoL in SCI so that potentially modifiable factors can be identified and specific interventions to improve HRQoL can be assessed.

Previous reports on HRQoL in SCI are limited by lack of accurate measurement of HRQoL, because instruments designed for the able-bodied or medical conditions unrelated to SCI were used. These instruments fail to consider the physical, social, and psychological burden that results from SCI. In addition, although a few previous reports have considered sociodemographics, severity of injury, time since injury, selected psychosocial variables, and ambulatory mode as factors that may be independently associated with HRQoL, most studies have not assessed these factors together. Respiratory dysfunction is common in SCI. However, the association of respiratory symptoms and pulmonary function with HRQoL has never been investigated.

There is no SCI-specific HRQoL instrument accepted for use in a clinical setting, nor has an instrument been administered systematically to a large number of persons with the goal of determining clinical factors potentially related to HRQoL in SCI. As part of a large cohort study assessing respiratory function in chronic SCI conducted at VA Boston, we used a 23-item SCI-specific HRQoL questionnaire (SCI QL-23), an HRQoL instrument developed by a Swedish group led by Sullivan. This instrument was specifically designed for people with SCI and includes questions addressing physical limitations (functioning domain); the perception of distress and depressive feelings (mood domain); perception of problems related to injury, such as bladder and bowel function (injury problem domain); and global quality of life. In this report, we assess the cross-sectional association of these HRQoL domains with objective measures of respiratory function and respiratory symptoms/comorbid illnesses. We also consider other important SCI characteristics such as severity of injury, time since injury, ambulatory mode, and sociodemographic factors.

MATERIALS AND METHODS

Patient Population

Since 1994, we have been assessing respiratory function in a chronic SCI cohort recruited from the VA Boston SCI Service and from the greater Boston area by advertisement, as described previously. A recruitment criterion of ≥1 yr post-SCI was selected to include participants who had survived acute injury and related complications. Participants requiring mechanical ventilation or having a tracheostomy were not assessed. At study entry, participants had to be free of acute illness. They completed a detailed general and respiratory health questionnaire and underwent pulmonary function testing and a neurological exam. From 1998 onwards, an assessment of HRQoL was introduced into our study, longitudinal follow-up began, and recruitment continued.

Participants (N = 391) completing a health assessment that included the HRQoL questionnaire between 1998 and June 2003 were included, and the first assessment was used in this study. We excluded from the analysis participants with a history of other neurological disorders such as polio, multiple sclerosis, and stroke (N = 25), one participant with a tracheostomy, and those with missing data on HRQoL (N = 9). The final dataset for analysis included 356 participants (262 [74%] veterans and 94 nonveterans). One person tested at 0.9 yrs after injury was retained in the cohort because there was no a priori basis to exclude him. The study was approved by the institutional review boards at VA Boston Healthcare System, Brigham and Women’s Hospital, and Harvard Medical School, and informed consent was obtained from each participant.

HRQoL Questionnaire

Each participant was asked to self-complete the SCI QL-23 at the time of pulmonary function testing. Few participants completed the questionnaire over the phone (N = 33; 9.3%). The phone tests were conducted if not enough time was available after completion of pulmonary function testing or if a participant was unable to travel. The SCI QL-23 was designed in a comprehensive Swedish research program and is reported to have high reliability and validity according to standard psychometric methodology. Cronbach’s alpha, a test for reliability, was 0.85–0.86 on various domains of the questionnaire (coefficients above 0.70 are considered satisfactory internal consistency). The SCI QL-23 reached content validity according to a conceptual model that includes condition-
specific aspects as well as generic aspects of physical and psychosocial functioning and well-being, and overall HRQoL. Unidimensionality of the three composite scales of SCI QL-23 was tested using principal component factor analysis, where factor loadings above 0.40 were considered. The clinical validity was checked through correlation analyses in subgroups representing individuals with tetraplegia and paraplegia, and individuals with complete and incomplete lesion, respectively. Another previously published cross-national comparison also has supported the validity of this instrument.29 We chose the SCI QL-23 questionnaire because it is SCI specific, short and easy to complete, even in a clinical setting, and yet comprehensive in eliciting possible items that are important to measure HRQoL in SCI.

The SCI QL-23 has four domains: functioning (ten items assessing physical and social limitations), mood (six items concerning distress and depressive feelings), SCI-related problems (six items regarding difficulty with loss of independence and other issues relating to injury, such as bladder and bowel function), and global HRQoL (assessed using a visual analog scale). HRQoL scores are generated individually for each of the four domains comprising three composite scale scores and one single-item score. Detailed information on response options, scoring instructions, handling of missing responses, and transformation to a 0–100 scale score can be obtained from the SCI QL-23 manual.30

Neurological Exam, Stature, and Weight

The assessment of motor level and completeness of injury was based on American Spinal Injury Association (ASIA) guidelines.31 Level and completeness of injury were determined by examination in all but two participants, whose level and completeness were determined by medical record review. Participants were a priori grouped into one of seven motor injury level and severity groups. These groups included motor complete SCI (cervical, high thoracic [T1–T6], and other lower levels). Participants with motor incomplete SCI (ASIA C: the majority of key muscles below the neurological level grade <3/5; or ASIA D: most muscles grade ≥3/5) were grouped into cervical C, other C, cervical D, and other D. Participants (n = 49) with motor complete SCI but evidence of some preservation of neurologic function below the neurologic level (more than two neurologic levels) were grouped with ASIA C participants. Weight and height were either self-reported or measured using previously described methodology.32

Health Questionnaire

The health questionnaire was based on the American Thoracic Society respiratory health questionnaire (ATS DLD-78)32 with supplemental questions. These supplemental questions elicit information on concurrent medical conditions such as heart disease and diabetes, and current and past employment. To assess mobility, participants were asked, “How you usually get around?” and were given the following options: motorized wheelchair more than half the time, hand-propelled wheelchair more than half the time, walk with aid more than half the time, or walk without assistance more than half the time. Using a series of structured questions, participants were asked to report activities of daily living that resulted in breathlessness. Participants were then asked to respond “yes” or “no” to the following questions: (A) Are you usually too breathless to leave the house or breathless while dressing or undressing? (B) Are you usually breathless while talking for more than a few minutes? (C) Are you usually breathless while eating? If the answer to one or more of the above questions was “yes,” the participant was classified as having dyspnea during activities of daily living. Chronic cough was defined as cough on most days for three consecutive months of the year, and chronic phlegm was defined similarly. Persistent wheeze was defined as wheeze reported on most days or nights, or with a cold and occasionally apart from colds.

Pulmonary Function Tests

Spirometry was based on American Thoracic Society standards34 modified for use in SCI, as described previously.35,36 Testing was done using a 10-liter water-seal spirometer, except in 31 participants where a water-seal portable spirometer was used (DSII or Survey III, Collins Pulmonary Diagnostics, Ferraris Respiratory, Louisville, CO). Predicted values for forced expiratory volume in 1 sec (FEV1) and forced vital capacity (FVC) were calculated using Hankinson et al.’s37 equations for Caucasians and African Americans. Some participants (n = 42) had missing information on FEV1 and FVC because health and HRQoL assessments were made on the phone for participants unable to travel to our study center (n = 20), because spirometry was not performed in 17 participants, or because predicted equations were not available for participants of other races (n = 5). In the 314 remaining participants, the best FVC and FEV1 were reported. Of these, 278 (86.1%) had at least three acceptable expiratory efforts with the best values of FEV1 and FVC, each within 200 ml; 11 (3.1%) were able to produce at least two acceptable and reproducible values of FEV1 and FVC; and two (0.6%) participants were only able to perform one acceptable effort. The remaining participants had two or three efforts that were acceptable but not reproducible.
Statistical Analysis

We examined cross-sectional determinants of each of the four domains of HRQoL using generalized linear models. Variables significant at the 0.10 level were subsequently assessed in multivariate models. Then, variables significant at the 0.05 level for any of the domains were included in the final model of baseline factors. The variables assessed for baseline model were sociodemographic characteristics including age, marital status, gender, race, years of schooling, and employment status; personal habits such as smoking status, and body mass index; SCI characteristics such as motor level and completeness of injury, years since injury, and age at injury; and mobility.

After the baseline model was determined, we assessed respiratory symptoms, comorbid medical conditions, and pulmonary function as possible predictors of HRQoL. A separate regression model was constructed for each of the symptom and pulmonary function variables by individually including each variable with the baseline model for each HRQoL domain. A sensitivity analysis was conducted by excluding participants who completed the questionnaires by telephone from final regression models. We performed statistical analysis using SAS for UNIX (version 9.0, SAS Institute Inc., Cary, NC).

RESULTS

Our cohort was well represented by participants of all age groups and varying severities of SCI (Table 1). A majority of participants were Caucasian (92.7%) and male (94.1%). The median time since injury was 18.2 yrs (range: 0.9, 57.9). Few participants reported physician-diagnosed respiratory illnesses (Table 2). However, the prevalence of chronic respiratory symptoms and heart disease treated in the last 10 yrs was more common.

In the baseline multivariate model, factors significantly associated with HRQoL included age, employment status, motor level and completeness of injury, and mobility (Table 3). Although age was significant only at the 0.10 level, it was retained a priori in the models because of its clinical significance. Interaction terms between variables in the baseline model were not statistically significant. The results suggest that older participants had a lower HRQoL compared with younger age groups on the functioning domain \( P < 0.05 \). Conversely, HRQoL on the injury problem scale was higher in participants greater than 60 yrs of age, especially in the 61- to 70-yr age group, compared with participants in the younger age groups. HRQoL on the functioning domain was lowest in the cervical motor complete SCI group compared with other complete SCI groups, and it also was lower in the cervical C SCI group than in the other C group. Similar findings were noted for mood state but not for perception of injury problems or global HRQoL.
domains. Each of the four HRQoL domains was significantly related to mobility.

Adjusting for covariates in the baseline multivariate model, participants with chronic cough, chronic phlegm, persistent wheezing, and dyspnea while talking, eating, or dressing had a significantly lower HRQoL on most domains (Table 4). The results were similar after further controlling for smoking status (current, ex, and never) in the multivariate models (data not shown). Mean values on the functioning domain of HRQoL were significantly lower (by 11 points; \( P < 0.01 \)) for participants who, in the year before questionnaire completion, had chest illnesses that had kept one at home, in bed, or out of work (n = 109; data not shown). Other comorbid illnesses, including physician-diagnosed asthma, chronic obstructive pulmonary disease, and heart disease treated in the last 10 yrs, were not significant predictors of HRQoL, possibly because fewer participants had these illnesses.

Adjusting for variables in the baseline multivariate model, percent predicted FEV\(_1\) and FVC (in quartiles) were significantly associated with HRQoL (Table 5). As percent predicted FEV\(_1\) and FVC (in quartiles) decreased, there was a decline in HRQoL. The results were similar after further adjusting for smoking status in the multivariate model (data not shown).

The results for all multivariate models were similar when participants who completed health and HRQoL questionnaires on the phone were excluded from the analysis (data not shown).

**DISCUSSION**

Previous studies have provided little guidance in identifying factors that are independently associated with HRQoL in SCI. In a large cross-sectional cohort of participants with SCI, we examined factors associated with HRQoL using an SCI-specific instrument. These factors included age, employment status, motor level and completeness of injury, and mobility. Respiratory symptoms and pulmonary function, the role of which has not been previously reported, were also associated with HRQoL in SCI.

The importance of improving HRQoL as one of the primary goals in SCI rehabilitation, as well as lack of appropriate assessment of HRQoL and factors influencing it, is well known and has been described in comprehensive reviews by Hallin et al.,\(^{38}\) Tulsky and Rosenthal,\(^{5}\) Hammell,\(^{39}\) and Wood-Dauphinee et al.\(^{6}\) The assessment of factors associated with HRQoL in SCI is a two-step process, neither of which have been adequately addressed in previous studies: first step is accurate measurement of HRQoL; second is to subsequently determine which factors independently influence HRQoL and consider them together in multivariate analysis. HRQoL is a construct that is specific to a particular disease, population, and treatment.\(^{5}\) Therefore, traditional generic instruments meant for use in the able bodied cannot be solely used to assess physical and social limitations, perceptions regarding loss of independence, or other injury-specific issues that constitute HRQoL in individuals with SCI. In contrast to diseases in the able bodied, where most investigators have recognized the importance of using disease-specific questionnaires to assess HRQoL,\(^{40}\) only two SCI-specific questionnaires have been developed.\(^{41}\) After HRQoL is accurately measured, factors that are potential determinants of HRQoL in a chronic disability such as SCI need to be considered and adjusted for. Among these possible determinants, we considered sociodemographics (age, marital status, education, and employment), injury-related factors (level and completeness of injury, years since injury, and age at injury), respiratory health (lung function, symptoms, and illnesses), other comorbidities, and functional status (the ability to get around, or mobility).

It is well established that the higher the level and the more complete the injury, the more likely it is that there will be loss of muscle function and

### TABLE 2 Medical comorbidities and pulmonary function in participants with spinal cord injury

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>(n (%))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medical comorbidities</td>
<td>(n = 356)‡</td>
</tr>
<tr>
<td>Physician-diagnosed chronic obstructive pulmonary disease</td>
<td>30 (8.4)</td>
</tr>
<tr>
<td>Physician-diagnosed asthma</td>
<td>33 (9.3)</td>
</tr>
<tr>
<td>Heart disease treated in last 10 yrs</td>
<td>32 (9.0)</td>
</tr>
<tr>
<td>Chronic cough</td>
<td>61 (17.1)</td>
</tr>
<tr>
<td>Chronic phlegm</td>
<td>73 (20.5)</td>
</tr>
<tr>
<td>Persistent wheeze</td>
<td>72 (20.2)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pulmonary function ((n = 314))</th>
<th>Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quartile 1</td>
<td>47.6 (10.4)</td>
</tr>
<tr>
<td>Quartile 2</td>
<td>70.3 (4.9)</td>
</tr>
<tr>
<td>Quartile 3</td>
<td>84.3 (3.4)</td>
</tr>
<tr>
<td>Quartile 4</td>
<td>99.9 (8.2)</td>
</tr>
<tr>
<td>Quartile 1</td>
<td>48.4 (10.9)</td>
</tr>
<tr>
<td>Quartile 2</td>
<td>70.2 (4.2)</td>
</tr>
<tr>
<td>Quartile 3</td>
<td>83.4 (3.5)</td>
</tr>
<tr>
<td>Quartile 4</td>
<td>99.2 (8.2)</td>
</tr>
</tbody>
</table>

FEV\(_1\), forced expiratory volume in 1 sec; FVC, forced vital capacity.
‡ Each row represents the number (%) of participants with a given comorbidity. Therefore, participants may be listed in several rows.
* Missing values for FEV\(_1\) and FVC = 42.
strength, and functional disability in SCI. However, previous studies have reported conflicting results on the association between level and completeness of injury and HRQoL. Some studies have reported a significant association between higher-level and more complete injury and a lower HRQoL, whereas others have found no such association. In our study, HRQoL for the functioning domain was lowest in participants with the most severe SCI (cervical motor complete SCI) compared with others with complete SCI, and it was also lower in higher levels of incomplete injury (cervical C) compared with other incomplete injuries. These findings suggest that the SCI QL-23 questionnaire is sensitive to clinically relevant functional differences and severities of SCI and, therefore, clinically valid. However, further work is needed to demonstrate the overall validity of the SCI QL-23 questionnaire, and its validity needs to be demonstrated in different SCI study populations. Similar findings were noted for the mood state domain but not for perception of injury problems or global HRQoL domains.

This might be because the functioning domain of SCI QL-23, which, by design, measures HRQoL on the basis of perceived functional status is

<table>
<thead>
<tr>
<th>TABLE 3 Adjusted mean scores for factors associated with health-related quality of life in 356 participants with spinal cord injury</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline Multivariate model</strong></td>
</tr>
<tr>
<td>Covariate</td>
</tr>
<tr>
<td>Age, yrs**</td>
</tr>
<tr>
<td>≤40</td>
</tr>
<tr>
<td>41–50</td>
</tr>
<tr>
<td>51–60</td>
</tr>
<tr>
<td>61–70</td>
</tr>
<tr>
<td>&gt;70</td>
</tr>
<tr>
<td>Employment status†</td>
</tr>
<tr>
<td>Full-time job</td>
</tr>
<tr>
<td>Part-time job</td>
</tr>
<tr>
<td>Not working because of disability/illness</td>
</tr>
<tr>
<td>Retired/unemployed/student</td>
</tr>
<tr>
<td>Motor level and severity of injury†</td>
</tr>
<tr>
<td>Motor complete</td>
</tr>
<tr>
<td>Cervical</td>
</tr>
<tr>
<td>High thoracic (T1–T6)</td>
</tr>
<tr>
<td>Other (low thoracic [T7–T12], lumbar, sacral)</td>
</tr>
<tr>
<td>Motor incomplete</td>
</tr>
<tr>
<td>Cervical C</td>
</tr>
<tr>
<td>Other C (thoracic, lumbar, sacral)</td>
</tr>
<tr>
<td>Cervical D</td>
</tr>
<tr>
<td>Other D (thoracic, lumbar, sacral)</td>
</tr>
<tr>
<td>Mobility (more than half the time)‡</td>
</tr>
<tr>
<td>Using motorized wheelchair</td>
</tr>
<tr>
<td>Using hand propelled wheelchair</td>
</tr>
<tr>
<td>Walk with crutches or cane</td>
</tr>
<tr>
<td>Walk without assistance</td>
</tr>
</tbody>
</table>

SE, standard error.
* Variable not significant at the 0.05 level; † P < 0.05 for all reported domains.
** P = 0.09 for functioning domain and P = 0.07 for injury problems domain for age.
† P < 0.05 for functioning domain and P = 0.09 for mood state domain for motor level and severity.
R² for baseline model: functioning = 44.0%; mood state = 14.9%; injury problems = 13.5%; global = 14.4%.
The higher the score on the functioning, mood state, and injury problems scales, the lower is the HRQoL on the respective scale. Higher scores on the global HRQoL scale represent a higher HRQoL.
most sensitive to changes with varying severities of SCI.

Older participants in our study had a lower HRQoL than younger age groups on the functioning domain. This is consistent with the results of some, but not all, previous cross-sectional studies. However, on the injury problem scale, HRQoL was better in older participants than in younger age-groups. It has been hypothesized that individuals with SCI adapt to their injury as they age, thereby leading to a better HRQoL on the injury problem scale. A longitudinal study is required to assess these observations further because it is likely that a cross-sectional study might underestimate the effect of aging on HRQoL with “healthier” participants surviving to be in our cohort. Those working in full- or part-time jobs had a better HRQoL on the functioning, mood, and global HRQoL domains. This relationship is similar to that described in some previous studies.

### TABLE 4 Adjusted mean scores for medical comorbidities associated with health-related quality of life in 356 participants with spinal cord injury

<table>
<thead>
<tr>
<th>HRQoL Domains (Scale 0–100)</th>
<th>n</th>
<th>Functioning</th>
<th>Mood State</th>
<th>Injury Problems</th>
<th>Global</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chronic cough†</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>61</td>
<td>33 ± 3</td>
<td>25 ± 2</td>
<td>*</td>
<td>71 ± 3</td>
</tr>
<tr>
<td>No</td>
<td>295</td>
<td>27 ± 2</td>
<td>16 ± 1</td>
<td>*</td>
<td>79 ± 2</td>
</tr>
<tr>
<td>Chronic phlegm†</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>73</td>
<td>34 ± 3</td>
<td>25 ± 2</td>
<td>33 ± 2</td>
<td>73 ± 3</td>
</tr>
<tr>
<td>No</td>
<td>283</td>
<td>26 ± 2</td>
<td>16 ± 1</td>
<td>29 ± 1</td>
<td>79 ± 2</td>
</tr>
<tr>
<td>Persistent wheeze‡</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>72</td>
<td>*</td>
<td>23 ± 2</td>
<td>*</td>
<td>72 ± 3</td>
</tr>
<tr>
<td>No</td>
<td>284</td>
<td>*</td>
<td>16 ± 1</td>
<td>*</td>
<td>79 ± 2</td>
</tr>
<tr>
<td>Dyspnea while talking, eating, or dressing‡</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>34</td>
<td>37 ± 4</td>
<td>*</td>
<td>*</td>
<td>69 ± 4</td>
</tr>
<tr>
<td>No</td>
<td>322</td>
<td>27 ± 2</td>
<td>*</td>
<td>*</td>
<td>78 ± 2</td>
</tr>
</tbody>
</table>

a Controlling for variables in the baseline multivariate model (Table 3).
b Each variable was added in separate regression models to the baseline multivariate model.
* Variable not significant at the 0.05 level.
† P ≤ 0.01 for all reported domains, except for functioning domain of chronic cough, where P = 0.06.
‡ P < 0.05 for all four domains.

Range of variance (R²) explained by each regression model: functioning = 44.5–45.2%; mood state = 17.5–20.1%; injury problems = 14.7%; global = 15.9–16.2%.

The higher the score on the functioning, mood state, and injury problems scales, the lower the HRQoL is on the respective scale. Higher scores on the global HRQoL scale represent higher HRQoL.

### TABLE 5 Adjusted mean scores for association of FEV₁ and FVC with health-related quality of life in 314 participants with spinal cord injury

<table>
<thead>
<tr>
<th>HRQoL Domain* (Scale 0–100)</th>
<th>Quartile 1 (Lowest percent predicted FEV₁ or FVC)</th>
<th>Quartile 2</th>
<th>Quartile 3</th>
<th>Quartile 4 (Highest percent predicted FEV₁ or FVC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent predicted FEV₁ †, quartile 1</td>
<td>35 ± 3</td>
<td>28 ± 3</td>
<td>24 ± 3</td>
<td>23 ± 3</td>
</tr>
<tr>
<td>Percent predicted FVC †, quartile 1</td>
<td>35 ± 3</td>
<td>28 ± 3</td>
<td>22 ± 3</td>
<td>24 ± 3</td>
</tr>
</tbody>
</table>

FEV₁, forced expiratory volume in 1 sec; FVC, forced vital capacity.
a Controlling for variables in the baseline multivariate model (Table 3).
b Each variable was added in separate regression models to the baseline multivariate model.
† P ≤ 0.01; * mood state, injury problems, and global HRQoL domains were not significant.

The variance (R²) explained by the regression model with FEV₁ = 47.2% and with FVC = 47.3%.
A higher score on the functioning scale reflects a lower HRQoL.
Participants using crutches or canes, or those using hand-propelled or motorized wheelchairs to get around, had a significantly lower HRQoL than those getting around without an assistive device. Hence, a higher HRQoL was related with the ability to get around independently (as in participants walking without assistance). We also found that those who usually walked with crutches or canes had a similar or lower HRQoL on most domains compared with those using hand-propelled or motorized wheelchairs. Because most participants using crutches or canes had lower SCI levels and incomplete injuries, we expected them to have a significantly higher HRQoL than those using motorized or hand-propelled wheelchairs (who are weaker and have higher levels and more complete injuries). It is possible that participants using ambulatory aids reported a lower HRQoL than would have been expected because of the greater effort and energy costs associated with using crutches or canes compared with hand-propelled or motorized wheelchairs. These cross-sectional results suggest that improving mobility or the ability to get around in an energy-efficient manner may improve HRQoL in SCI.

The increased prevalence of dyspnea in SCI has been reported previously. We also have reported previously that participants using crutches or canes have increased prevalence of dyspnea when performing activities of daily living compared with those using hand-propelled wheelchairs. Our results from this previous study and the current study imply that participants using crutches or canes to get around have not only a lower than expected HRQoL but also a higher than expected prevalence of dyspnea (which was associated with lower HRQoL in our study). Therefore, improving mobility may help reduce the prevalence of dyspnea and improve HRQoL.

Although respiratory dysfunction is common in SCI, the association of respiratory symptoms and comorbid illnesses with HRQoL has not been previously investigated. Because HRQoL was lower in participants reporting chronic respiratory symptoms and chest illnesses, therapy directed toward the recognition and treatment of these conditions may improve HRQoL in SCI. Our study is the first to report on the association of pulmonary function with HRQoL in SCI. People with reduced percent predicted FEV₁ and FVC had lower HRQoL on the functioning domain. Similar results have been reported from other able-bodied cohorts of participants with chronic obstructive lung disease. It is possible that interventions that lead to improvements in FEV₁ and FVC or that lead to improvements in functional ability may improve HRQoL in SCI. Because we performed a cross-sectional analysis, it is not possible to assess whether reduced pulmonary function led to lower HRQoL on the functioning domain or whether greater functional disability (leading to lower HRQoL on the functioning domain) resulted in decreased pulmonary function. A longitudinal study would be needed to assess the direction of a causal relationship between functioning domain of HRQoL and pulmonary function.

Clinicians not familiar with HRQoL concepts may find it difficult to interpret a meaningful change in HRQoL score and its corresponding clinical relevance. Our cross-sectional results can help guide the clinician about the health implications of a change in SCI QL-23 score. For instance, a seven-point improvement in HRQoL score (from 35 to 28) on the functioning domain was associated with an increase in mean percent predicted FEV₁ from 47.6% (quartile 1) to 70.3% (quartile 2). Similarly, a ten-point improvement on the functioning domain (difference of means between 37 and 27 points) was associated with the absence of dyspnea during activities of daily living. On average, an eight-point improvement on the mood scale was associated with participants being able to work full time compared with those not working because of disability or illness. These results can help one put into perspective the clinical relevance of changes in HRQoL scores when future interventions and their impact on HRQoL are assessed.

Although we identified factors significantly associated with HRQoL, overall, the variance explained by the regression models was modest and varied by domain (range, 13.5–44%; Tables 3-5), suggesting that there are additional factors associated with HRQoL in this cohort. For example, our study did not include a detailed assessment of psychosocial and other social support factors that may influence HRQoL in SCI. Another limitation is the participation of few females and minorities in our study. Hence, factors specific to these populations may not have been accounted for in our analysis. Although disease specific instruments more accurately measure HRQoL than generic instruments (as described earlier), this approach may lead to colinearity between the independent and dependent variables when studying factors associated with HRQoL. For example, questions used to define HRQoL in SCI may include descriptions of specific functional and psychological limitations that also are used as independent (predictive) variables in a regression model.

In summary, our study identified age, employment status, motor level and completeness of injury, and mobility as independent factors associated with HRQoL in SCI. The association of respiratory symptoms and illnesses, and decreased FEV₁ and FVC, with a lower HRQoL not only demonstrates previously unreported factors influencing
HRQoL; it also shows the sensitivity and clinical validity of the SCI QL-23 instrument. Our study reports several modifiable factors that, if addressed by a clinician, could potentially lead to HRQoL improvement in individuals with SCI. Our results also can help the clinician understand the health implications of a change in HRQoL scores when future interventions to improve HRQoL, as measured by SCI QL-23, are reported.

ACKNOWLEDGMENTS

We gratefully acknowledge Honghu Guan for assistance with statistical programming, and Sarah Curran, Sarah McGee, Kara Zayac, and Kirby Mattthes for helping with data collection and management of our study. As always, we thank physicians and staff of the Spinal Cord Injury unit at Veterans Affairs hospital in West Roxbury for their assistance, without which this study would not have been possible. Finally, we acknowledge the enthusiastic participation of all participants in our cohort.

REFERENCES

23. Jain NB, Brown R, Tun CG, Gagnon D, Garshick E: Determinants of forced expiratory volume in 1 second (FEV1), forced vital capacity (FVC), and FEV1/FVC in chronic spinal cord injury. *Arch Phys Med Rehabil* 2006;87:1327–33


Dry Needling to a Key Myofascial Trigger Point May Reduce the Irritability of Satellite MTrPs

ABSTRACT


Objective: To investigate the changes in pressure pain threshold of the secondary (satellite) myofascial trigger points (MTrPs) after dry needling of a primary (key) active MTrP.

Design: Single blinded within-subject design, with the same subjects serving as their own controls (randomized). Fourteen patients with bilateral shoulder pain and active MTrPs in bilateral infraspinatus muscles were involved. An MTrP in the infraspinatus muscle on a randomly selected side was dry needled, and the MTrP on the contralateral side was not (control). Shoulder pain intensity, range of motion (ROM) of shoulder internal rotation, and pressure pain threshold of the MTrPs in the infraspinatus, anterior deltoid, and extensor carpi radialis longus muscles were measured in both sides before and immediately after dry needling.

Results: Both active and passive ROM of shoulder internal rotation, and the pressure pain threshold of MTrPs on the treated side, were significantly increased (P < 0.01), and the pain intensity of the treated shoulder was significantly reduced (P < 0.001) after dry needling. However, there were no significant changes in all parameters in the control (untreated) side. Percent changes in the data after needling were also analyzed. For every parameter, the percent change was significantly higher in the treated side than in the control side.

Conclusions: This study provides evidence that dry needle-evoked inactivation of a primary (key) MTrP inhibits the activity in satellite MTrPs situated in its zone of pain referral. This supports the concept that activity in a primary MTrP leads to the development of activity in satellite MTrPs and the suggested spinal cord mechanism responsible for this phenomenon.

Key Words: Dry Needling, Myofascial Trigger Point, Referred Pain
A myofascial trigger point (MTrP) has been defined as a hyperirritable (hypersensitive) spot in a taut band of skeletal muscle fibers. It has been shown that a latent MTrP (i.e., one that is exquisitely tender but not a source of pain) can be identified in most skeletal muscles. A latent MTrP can be activated to become an active MTrP, which is painful and very tender. In clinical observations, when an active MTrP is suppressed, it is still tender but not painful, and it becomes a latent MTrP. The latent MTrP may be activated to become an active one secondary to or associated with various pathologic conditions. After appropriate treatment or control of this condition, the activated MTrP can be suppressed to become inactive. The MTrPs do not disappear; rather, they change from active to latent. There are two important characteristics of an active MTrP. One is pain referred from a distant site, the referred pain (ReP). The other is a local twitch response (LTR), which is a brisk contraction of a group of muscle fibers in a taut band, in response to a rapid, brief mechanical stimulation at an active MTrP site.

Recent studies have helped make the pathophysiology of MTrP much clearer. In an animal model, Hong and Torigoe have observed that when pressure was applied at a hyperirritable site in the skeletal muscle of a rabbit, the animal showed evidence of severe discomfort. The latter, however, did not occur when similar pressure was applied at a nonsensitive site. In addition, at a hyperirritable site of this type, many LTRs could be elicited. On the basis of this animal model, there is reason to believe that an LTR is a spinal cord-mediated reflex. There are a large number of sensitive loci in the region of an MTrP. When a sensitive locus is mechanically stimulated by a needle tip (high-pressure stimulation), an LTR can be elicited. This locus has been defined as an LTR locus. An LTR locus has been shown to be one in which there are numerous nociceptive nerve endings. These LTR loci are most frequently found in a muscle’s endplate zone. Endplate noise has been recorded much more frequently at MTrP sites than at any other parts of skeletal muscle in both human and animal studies. Simons and colleagues have suggested that the taut bands found at MTrP sites may develop as a result of excessive acetylcholine leakage from dysfunctional motor nerve terminals.

In humans, stimulation of the LTR locus can elicit pain (low-pressure stimulation), ReP (moderate-pressure stimulation), and LTR (high-pressure stimulation). However, when the MTrP is hyperirritable, even low-pressure stimulation can elicit ReP and LTR. It has been suggested that the degree of irritability is proportionate to the number of LTR loci (sensitized nociceptors) in the MTrP region. It has been demonstrated that dry needling of MTrP is effective for pain relief. In a human study on humans, Hong has demonstrated that either injection of a local anesthetic agent into or dry needling carried out at an MTrP site are similarly effective in alleviating MTrP pain, as long as LTRs are elicited while carrying out either of these procedures. The pain-relieving mechanisms brought into action, however, are still unclear, even though many authors have stressed the importance of eliciting LTRs while carrying them out to obtain immediate and complete pain relief.

In our clinical practice, we have observed the phenomena of interactions among primary (key) and satellite MTrPs situated in the zone of pain referral. A patient who has multiple MTrPs, if a certain MTrP (primary or key MTrP) has been inactivated by means of needle stimulation, other MTrPs (secondary or satellite MTrPs) also can be suppressed. To our knowledge, no clinical trial has previously been carried out to provide objective evidence of this phenomenon. The purpose of this study, therefore, was to do this by means of a single blinded study in which active bilateral infraspinatus MTrPs would be dry needled in a group of patients suffering from bilateral shoulder pain.

**MATERIALS AND METHODS**

**General Design**

Patients who had bilateral shoulder pain with active MTrPs in the infraspinatus muscles on both sides were recruited for this study. Each patient received treatment with dry needling of the MTrP in the infraspinatus muscle in a randomly selected side, but no treatment of the MTrP on the control side. The range of motion (ROM) of shoulder internal rotation, the pain intensity in the shoulder (including MTrPs in infraspinatus), and the pressure pain thresholds of MTrPs in the infraspinatus, anterior deltoid, and extensor carpi radialis longus muscles were assessed in both sides (experiment = treated side; control = untreated side) before and immediately after dry needling. Each subject was served for both control and experimental groups, with no treatment on one side and dry needling on the other side. In this way, the homogeneity (similar tissues in one human body) of the samples could be improved. Normalization of data (percentages of differences between the pre- and posttreatment data) for statistical analysis was also performed to eliminate the bias from the differences in the pretreatment data between two groups.
Subjects

For this study, 14 patients (eight males and six females, average age: 60.2 ± 13.2 yrs) with bilateral shoulder pain were selected from a pain-control clinic of a university hospital for this study. The causes of shoulder pain included fibromyalgia (three males and three females), subacromial bursitis (four males and three females), and rotator cuff tendonopathy (one male). Each subject signed the consent form, which had been approved by our university's institutional review board. The inclusion/exclusion criteria of selection included:

1. Having bilateral shoulder pain without treatment other than oral medication for at least 3 mos.
2. Having active MTrPs in the infraspinatus muscles in both sides.
3. Having no contraindication for needling of infraspinatus muscle, such as local infection, serious medical problems, recent multiple trauma, or pregnancy with threatened abortion.
4. Having no condition such as substance abuse (including alcohol) that might interfere with the assessment of pain or pain threshold.
5. Having had no previous surgery to the neck or upper limb.
6. Having no significant differences in clinical presentation (such as pain intensity) between two sides.

Assessment of ROM

For the measurement of ROM, each patient was placed in a comfortable sitting position with the shoulder abducted at 90 degrees and elbow flexed at 90 degrees. Then, the patient was asked to move his or her hand forward and downward (internal rotation of shoulder) as far as possible. The arc of this movement was measured with a large goniometer. After this, the passive ROM was measured in both shoulders by pushing the internal rotation movement further to the endpoint (limited either by pain or by tightness). Active and passive ROMs were measured in bilateral shoulders before and immediately after dry needling of the infraspinatus MTrP.

Assessment of Pain Intensity

The subject was requested to describe the pain intensity in both shoulders before and immediately after dry needling. Pain intensity was assessed by means of the use of a visual analog pain scale, which is a card with an uncalibrated scale ranging from zero to ten on one side (with zero representing no pain and ten representing the worst imaginable pain) and a corresponding 10-cm ruler on the other side (with each centimeter representing one pain level). The patient subjectively estimated his/her pain level by moving the pointing device along the uncalibrated scale between zero to ten. Then, the exact value of pain intensity could be obtained by referring the uncalibrated scale to the ruler on the back side.

Assessment of Pressure Pain Threshold

The MTrPs of bilateral infraspinatus, anterior deltoid, and extensor carpi radialis longus muscles were identified by finger palpation of the hyperirritable spots in taut bands, as described by Simons et al. The selection of these muscles was based on the fact that both anterior deltoid and extensor carpi radialis longus muscles are in the ReP zone of the infraspinatus MTrP.

These MTrPs were marked for the measurements of pressure pain threshold, so that the three consecutive measurements could be performed over the same site. The pressure pain thresholds of these MTrPs were measured by a well-trained assistant who was blinded to the side of treatment. A pressure-threshold algometer developed by Fischer was used for measuring this, because this algometer has proved to be both reliable and valid. The procedure of pain-threshold measurement recommended by Fischer was applied in this study. First, the procedure was explained clearly to the patient, who was then placed in a comfortable sitting position and encouraged to maintain complete relaxation. The algometer was applied on the marked site with the metal rod perpendicular to the surface of the skin. The pressure of compression was increased gradually at a speed of approximately 1 kg/sec. The subject was asked to say “yes” as soon as he or she began to feel pain or discomfort (for latent MTrPs) or an increase in pain intensity or discomfort (for active MTrPs). The compression was then stopped, and the subject was asked to remember this level of pain discomfort and to apply the same criterion for the subsequent measurements. The subject might demonstrate pain by pulling away or grimacing, indicating that the pain threshold had been exceeded. In such cases, the instruction was repeated and another measurement was taken to ensure that the “real” threshold was obtained. Three repetitive measurements at an interval of 30–60 secs were performed at each site. The average value of the three readings was used for data analysis. The pressure pain threshold is measured in kilograms per squared centimeter.

Procedure of Dry Needling

The MTrP dry needling procedure employed was similar to the MTrP injection described by Hong. The MTrP was located by palpating the
taut band and identifying the point of maximal tenderness. This was then firmly compressed by the index finger or middle finger of the nondominant hand to direct the placement of the needle tip while inserting the needle. A 5-ml syringe connected with a #25 hypodermic needle, 1.5 inches in length, was held by the dominant hand. The needle was inserted into the skin at a point above the taut band, approximately 1 cm from the MTrP region. After penetration of the needle into the subcutaneous layer, it was kept there and obliquely (about 45 degrees) directed to the MTrP region under the fingertip of the nondominant hand. Then, the needle was inserted rapidly into the MTrP region and withdrawn rapidly. With rapid movement of needle, an LTR can always be elicited if the needle tip encounters a sensitive locus (LTR locus). The reason for employing rapid needle movements is to provide high-pressure stimulation for eliciting LTRs and to avoid side movement of the needle that may side cut (stretch) the muscle fibers. The needle insertions were repeated to elicit as many LTRs as possible. Usually, 1–2 mins were required for the complete procedure in each MTrP region. As soon as the needle was pulled out of skin, the MTrP region and the open wound of the needle-insertion site were compressed firmly for at least 3 mins to prevent excessive bleeding.

Data Analysis

The mean and standard deviation of the values measured for ROM, pain intensity, and pressure pain threshold were calculated. The paired Student’s t test was used to assess the differences between the data before and after needling, and the differences in the data between the treated and untreated sides. The differences in ROM, pain intensity, and pain threshold after needling were further normalized as follows: percentage of changes = [(data after treatment − data before treatment)/data before treatment] × 100%. After data normalization, as described above, the differences in the changes of ROM, pain intensity, or pain threshold between the two sides (treated and untreated) were compared with paired t test. The confidence interval was set at 95% (P < 0.05). All data were analyzed using Statistical Package for the Social Sciences version 8.0 for Windows.

RESULTS

The changes in the investigated parameters are listed in Tables 1-3.

Increase in Mobility of Shoulder after Needling

As shown in Table 1, there were significant increases in both active and passive ROMs in the treated shoulder after dry needling (P < 0.01), and there were no significant changes in active or passive ROMs in the untreated shoulder. After normalization of data, the percent increases were significantly bigger in the treated side than in the untreated side (P < 0.01).

Reduced Pain Intensity

Subjectively, all patients had remarkably reduced pain of the shoulder in the treated side but only little (if any) pain relief in the untreated side. No subject had pain in the MTrPs of anterior deltoid and extensor carpi radialis longus muscles; they were latent MTrPs. After dry needling, the pain intensity was significantly reduced in the treated shoulder (P < 0.001) but not in the untreated shoulder (P > 0.005). Comparing the normalized data, the mean percent decrease in pain intensity was significantly higher in the treated shoulders than in the untreated ones (P < 0.001).

Increase in Pressure Pain Threshold of MTrPs

As shown in Table 3, after dry needling, there were significant increases in pressure pain threshold of MTrPs in the treated shoulder (P < 0.01) and there were no significant changes in the untreated shoulder. After normalization of data, the mean percent increase in pressure pain threshold was significantly higher in the treated side than in the untreated side (P < 0.01).

| TABLE 1 Active and passive range of motion of shoulder internal rotation (degrees) before and after dry needling |
|---------------------------------------------------------------|--------------------------|-------------------------------------------------|------------------|
| Active range of motion                                       | Before Needling          | After Needling                                 | % Changes        | P Values (Before vs. After) |
| Treated side                                                 | 47.5 ± 16.4              | 70.7 ± 16.5                                   | 55.1 ± 31.0 (%)  | <0.01                       |
| Untreated side                                               | 50.4 ± 13.7              | 54.3 ± 16.3                                   | 7.1 ± 8.8 (%)    | <0.1                        |
| P values (treated vs. untreated)                             |                          |                                                |                  |                             |
| Passive range of motion                                      | Treated side             | 51.8 ± 15.5                                   | 55.1 ± 28.3 (%)  | <0.01                       |
| Untreated side                                               | 52.5 ± 14.2              | 61.4 ± 18.2                                   | 16.6 ± 12.0 (%)  | >0.05                       |
| P values (treated vs. untreated)                             |                          |                                                |                  |                             |
TABLE 2 Pain intensity of shoulders before and after dry needling

<table>
<thead>
<tr>
<th></th>
<th>Before Needling</th>
<th>After Needling</th>
<th>% Changes (Normalized Data)</th>
<th>P Values (Before vs. After)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treated side</td>
<td>7.8 ± 1.2</td>
<td>2.8 ± 1.1</td>
<td>−64.8 ± 12.6 (%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Untreated side</td>
<td>7.7 ± 1.4</td>
<td>6.8 ± 1.3</td>
<td>−14.7 ± 7.8 (%)</td>
<td>&gt;0.05</td>
</tr>
<tr>
<td>P values (treated vs. untreated)</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

DISCUSSION

In this study, we have demonstrated that after dry needling of an MTrP in the infraspinatus muscle, the pressure pain thresholds of the MTrPs in the treated infraspinatus muscle and the ipsilateral anterior deltoid and extensor carpi radialis longus muscles (located in the ReP zone of the MTrP in the infraspinatus) were significantly increased, in addition to the significant improvement in the ROM and pain intensity of the treated shoulder.

This important finding further supports the concepts of key MTrP and satellite MTrPs and the possible spinal cord mechanism of this phenomenon, as explained below. We also have further confirmed the effectiveness of dry needling, as has been well demonstrated in the previous studies.1,24–30

On the basis of clinical as well as basic studies, Hong and colleagues6,7,11–13,21,29 have shown that there are multiple sensitive loci in an active MTrP region. A region’s irritability is probably proportionate to the number of sensitive loci (LTR loci) and sensitized nociceptors it contains. Furthermore, these sensitized nociceptors are capable of sending enough neural impulses to the spinal cord to induce central sensitization of some dorsal horn cells to which MTrPs in the referred zone project. This active MTrP is the key MTrP, and the sensitized MTrPs in the referred zone are the satellite MTrPs. In this way, the receptive fields of the key MTrP are expanded.36–39 Therefore, the pressure pain threshold of the satellite MTrPs are reduced because of central sensitization. When the irritability of a key MTrP is suppressed after appropriate treatment, the irritability of the satellite MTrPs in the referred zone can also be reduced because of the removal of central sensitization. This mechanism can explain the phenomenon observed in our current study.

Simons et al.1 and Travell and Simons2 have shown that the MTrP pain referral pattern for each

TABLE 3 Pressure pain threshold (kg/cm²) of myofascial trigger points before and after dry needling

<table>
<thead>
<tr>
<th></th>
<th>Before Needling</th>
<th>After Needling</th>
<th>% Changes (Normalized Data)</th>
<th>P Values (Before vs. After)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infraspinatus</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treated side</td>
<td>2.3 ± 0.5</td>
<td>4.1 ± 0.5</td>
<td>80.2 ± 30.7 (%)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Untreated side</td>
<td>2.5 ± 0.5</td>
<td>2.7 ± 0.5</td>
<td>11.3 ± 6.0 (%)</td>
<td>&gt;0.05</td>
</tr>
<tr>
<td>P values (treated vs. untreated)</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anterior deltoid</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treated side</td>
<td>3.5 ± 0.5</td>
<td>4.5 ± 0.4</td>
<td>30.8 ± 15.1 (%)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Untreated side</td>
<td>3.5 ± 0.5</td>
<td>3.6 ± 0.5</td>
<td>5.2 ± 4.4 (%)</td>
<td>&gt;0.05</td>
</tr>
<tr>
<td>P values (treated vs. untreated)</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extensor carpi radialis longus</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treated side</td>
<td>4.2 ± 0.5</td>
<td>4.7 ± 0.4</td>
<td>18.2 ± 9.9 (%)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Untreated side</td>
<td>4.0 ± 0.5</td>
<td>4.1 ± 0.5</td>
<td>3.9 ± 2.1 (%)</td>
<td>&gt;0.1</td>
</tr>
<tr>
<td>P values (treated vs. untreated)</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
individual muscle is the same in everyone. This also is true in animal studies.46–49 In other words, the connection of dorsal horn neurons among the key MTrP and satellite MTrPs consistently follows a certain pattern for each individual muscle. The ReP patterns of MTrPs in some muscles are similar to the distribution of traditional Chinese acupuncture meridians.30 The mechanism responsible for the pain-relieving effects of dry needling (including acupuncture) an MTrP and of injecting a local anesthetic into it must be similar.21,29,30,40,41

In acupuncture therapy, remote effectiveness in pain control has been documented.42–45 Needling of a point in the first dorsal interosseous muscles (Ho-Ku point) can effectively control headache or toothache. This phenomenon also may be related to the spinal cord mechanism described above. The effects of acupuncture also may spread to contralateral side. In our study, we have observed a trend of changes in the contralateral side, although the changes are not statistically significant. Those changes could be attributed to repeated measures with the algometer. However, in the previous studies on pressure pain threshold, no such phenomena have been observed.20,31,32 It is very likely that needling to the MTrP has an expanding effect to the opposite side, as is sometimes seen in acupuncture therapy.

The importance of eliciting LTRs to suppress MTrP pain is still unclear. It was found that after several LTRs had been elicited by the needling of an MTrP, no more LTRs could be elicited from the same region,11 and the irritability of the MTrP could be suppressed.46 It seems that there are certain neural connections in the spinal dorsal horns that control the irritability of an MTrP. These neural connections in the spinal cord may play an important role for ReP36–39,47 and LTR.11–13,21,29,48 The whole unit of these neural connections has been defined as an MTrP circuit.49 A strong pressure stimulation to the nociceptors in the MTrP region could elicit an LTR and could probably provide very strong neural impulses to the MTrP circuit to break the vicious cycle so that the MTrP pain could be relieved.30,49

The limitations of this study include the small sample size and the lack of appropriate controls. There was no dry needling to the untreated side, and there was no comparable “sham” procedure. This raises the possibility of nonspecific findings, such as a placebo effect, associated with the supposedly beneficial findings, and not effects of the treatment per se. For a subjective measurement such as pain intensity, the placebo effect would be critical. The measurement of ROM is an objective assessment, and the measurement of pressure pain threshold can be considered semiobjective if it is performed appropriately.31–35 Furthermore, this study was not designed to examine the therapeutic effectiveness of dry needling. We are more concerned with the changes in pressure pain threshold. For this semiobjective measurement, the blinded design is more important than the placebo design. In a future study, we should increase our sample size and apply a sham acupuncture procedure to confirm the current findings.

CONCLUSION
Immediately after dry needling of MTrP in the infraspinatus muscle, the pressure pain thresholds of MTrPs in the treated infraspinatus muscle and the ipsilateral anterior deltoid and extensor carpi radialis longus muscles (located in the referred zone of MTrP in the infraspinatus) can also be suppressed, in addition to the significant improvement in ROM and pain intensity of the treated shoulder. It is possible that, in some situations, inactivation of a key MTrP can suppress the irritability of its satellite MTrPs. These important findings further support the concepts of primary (key) MTrP and secondary (satellite) MTrPs, and the hypothetical spinal cord mechanism of this phenomenon.

REFERENCES
13. Hong CZ, Simons DG: Pathophysiologic and electrophysi-


May 2007
Measurement of Plantarflexor Spasticity in Traumatic Brain Injury
Correlational Study of Resistance Torque Compared with the Modified Ashworth Scale

ABSTRACT


Objectives: To examine the usefulness of a biomechanical measure, resistance torque (RT), in quantifying spasticity by comparing its use with a clinical scale, the modified Ashworth scale (MAS), and quantitative electrophysiological measures.

Design: This is a correlational study of spasticity measurements in 34 adults with traumatic brain injury and plantarflexor spasticity. Plantarflexor spasticity was measured in the seated position before and after cryotherapy using the MAS and also by strapping each subject’s foot and ankle to an apparatus that provided a ramp and hold stretch. The quantitative measures were (1) reflex threshold angle (RTA) calculated through electromyographic signals and joint angle traces, (2) H_{dorsiflexion}/H_{control} (Hdf/Hctrl) amplitude ratio obtained through reciprocal inhibition of the soleus H-reflex, (3) H_{vibration}/H_{control} (Hvib/Hctrl) ratio obtained through vibratory inhibition of the soleus H-reflex, and (4) RT calculated as the time integral of the torque graph between the starting and ending pulses of the stretch.

Results: Correlation coefficients between RT and MAS scores in both pre-ice (0.41) and post-ice trials (0.42) were fair (P < 0.001). The correlation coefficients between RT scores and RTA scores in both the pre-ice (0.66) and post-ice trials (0.75) were moderate (P ≤ 0.001).

Conclusion: RT is a measure of the cumulative torque during an imposed stretch. The MAS is a subjective measure of the cumulative resistance perceived by the clinician during an imposed stretch. RT seems to be a fair quantitative correlate of the MAS in assessing spasticity.

Key Words: Brain Injury, Spasticity, Reflex, Stretch, Biomechanics, Measurement
Spasticity in traumatic brain injury, stroke, and spinal cord injury has been known to interfere with function, limit independence, and produce secondary complications such as contractures. The most commonly used definition of spasticity was given by Lance in 1980: “Spasticity is a motor disorder characterized by velocity-dependent increase in tonic reflexes (muscle tone) with exaggerated tendon jerks, resulting from hyper excitability of the stretch reflex as one component of the upper motor neuron syndrome.” Even though spasticity has been well characterized over the years, its quantification by reliable and widely accepted standard methods remains as an unresolved problem in physical medicine and rehabilitation.

The effectiveness of therapeutic modalities used to treat spasticity has been evaluated using qualitative or semiquantitative clinical observations, such as tendon tap response, clonus, and Ashworth scale. Inherent limitations of these clinical scales may interfere with accurate assessment of spasticity that may then lead to poor evaluation and treatment. To study the effectiveness of therapeutic intervention, it is essential to first have an objective, quantifiable, reliable method of measuring spasticity. For example, reliable measurement of day-to-day spasticity variations for a particular joint of a given subject is likely to be extremely valuable in quantifying the effects of a therapeutic intervention such as a drug or surgical procedure.

The modified Ashworth scale (MAS) assigns grades to a manually determined resistance of muscle to passive stretching (Table 1). This scale offers qualitative information, but it has been reported to lack temporal and interexaminer reproducibility and to suffer from clustering effect in that most patients are grouped within the middle grades. Nevertheless, it has been and continues to be widely used in the measurement of spasticity, and it has been described by some authors as arguably the best available yardstick against which newer, more exact methods must be compared.

The MAS and other scales used to qualitatively assess spasticity correlate poorly with each other, suggesting that they assess different aspects of spasticity. The use of any single scale is likely to underrepresent the magnitude and severity of spasticity. A major shortcoming of qualitative scales such as the MAS is that the stimulus is not well controlled, and the output is difficult to quantify. An objective quantitative measure would achieve widespread acceptance only if its variations broadly paralleled an accepted “gold standard” clinical scale, but in the case of spasticity evaluation, there is no such universally accepted gold standard against which to compare putative systems for assessing spastic individuals. Important criteria that objective parameters must fulfill to gain everyday clinical acceptance are consistency, sensitivity, and ease of use. Spasticity is more difficult to characterize than recognize, and it is still more difficult to quantify. Therefore, there are currently no clear-cut winners in the quest of identifying an ideal quantitative measure of spasticity.

Proposed quantitative measures of spasticity typically involve electrophysiological, electromyographic, or biomechanical approaches and measure forces generated by muscles during movement or evaluate excitability of the stretch reflex loop, in an attempt to address underlying mechanisms of spasticity. However these quantitative measures have not gained wide acceptance, for a variety of reasons. There is a perception that quantitative measures do not correlate well with clinical measures. A partial explanation for this lack of correlation may be found in the argument by Priebe et al. that the various clinical measures assess different aspects of spasticity. Furthermore, some of the quantitative methods suggested are difficult to employ in a routine clinical setting, requiring expensive equipment, extensive setup, or both. But these measures provide meaningful metrics that are objective, eliminating complications associated with interrater reliability. Quantitative measurements also provide greater resolution than is available using standard clinical scales of spasticity.

Electrophysiological efforts to quantify spasticity have focused largely on the H-reflex and various mechanisms that modify its amplitude. For example, in an earlier study, of which this study is an extension, Allison and Abra-

### TABLE 1 The modified Ashworth scale (MAS) for assessing muscle spasticity

<table>
<thead>
<tr>
<th>Grade</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No increase in muscle tone</td>
</tr>
<tr>
<td>1</td>
<td>Slight increase in muscle tone, manifested by a catch and release or by minimal resistance at the end of range of motion when the affected part is moved</td>
</tr>
<tr>
<td>1+</td>
<td>Slight increase in muscle tone, manifested by a catch, followed by minimal resistance throughout the remainder (less than half) of the range of motion</td>
</tr>
<tr>
<td>2</td>
<td>More marked increase in muscle tone through most of the range of motion, but affected part easily moved</td>
</tr>
<tr>
<td>3</td>
<td>Considerable increase in muscle tone, passive movement difficult</td>
</tr>
<tr>
<td>4</td>
<td>Affected part rigid in flexion or extension</td>
</tr>
</tbody>
</table>

have described the use of two electrophysiological measures to quantify spasticity: the inhibition of the H-reflex amplitude attributable to reciprocal inhibition \( \text{H}_\text{dorsiflexion} / \text{H}_\text{control} \), and the inhibition of the H-reflex amplitude attributable to vibration of the antagonist \( \text{H}_\text{vib}/\text{H}_\text{ctrl} \). However, electrophysiological measurements are limited to local, circumscribed sampling of the potentials generated by the muscle and may, therefore, not be representative of the action of the total muscle. Electrophysiological measures of spasticity also require sophisticated equipment for recording and are generally regarded as impractical for use in clinical settings.

Biomechanical approaches can be classified into gravitational method, manual method, threshold angle, and controlled torque method, and these approaches have been studied extensively.\(^3\)\(^,\)\(^17\)\(^–\)\(^19\)

The controlled torque method involves instrumented measurements of the resistance to imposed movements of a limb joint and mimics the traditional clinical approach for spasticity evaluation in which a clinician estimates the resistance to a manually applied stretch of the joint. However, a mechanically imposed stretch is much more repeatable and consistent across trials and laboratories than a manually applied stretch. Recently, clinically available isokinetic dynamometers have been implemented to quantify spasticity in which the resistance to an imposed ramp and hold stretch of the joint is used as a metric of spasticity.\(^20\)

Katz et al.\(^5\) used reflex threshold angle (RTA), which they calculated as the angle during an imposed stretch at which there was a simultaneous increase in the slope of the torque graph and the onset of reflex electromyographic (EMG) activity and a measure of torque, which was calculated as an instantaneous value at a predetermined angle during a ramp and hold stretch. They report a significant correlation between RTA and MAS scores and a nonsignificant correlation between the other torque measure and MAS scores.

Firoozbakhsh et al.\(^17\) used maximum torque, sum of torque amplitudes, and slope of torque–velocity graphs during stretch, comparing these measures for normal and spastic subjects. They report a significant difference in the sum of torque amplitudes between normal and spastic subjects and higher slopes and sensitivity to rate of stretch in the spastic subjects compared with the normal subjects. Lehmann et al.\(^3\) applied a sinusoidal displacement to the ankles to measure torque in subjects with spinal cord injury–related spasticity. They report that total (elastic plus viscous) stiffness was four times that of normal subjects. However, including the above studies, there have been no published studies that have correlated torque measures with MAS scores in the clinical assessment of the severity or extent of plantarflexor spasticity.

By definition, spastic limbs demonstrate abnormal resistance to externally imposed joint movement. This resistance is augmented by increasing the angle of deflection and the rate at which the limb is moved. Torque is proportional to the amount of muscle force elicited by moving that limb through a specified angle. Torque as a measure of resistance to imposed movement is the result not only of the reflex response but also of the passive properties of the tissues.\(^3\) An ideal torque measure would have a significant correlation with the MAS, demonstrate change similar to it in response to treatment strategies such as cryotherapy,\(^21\) and have a strong theoretical basis behind its development.

Allison and Abraham\(^11\) used correlation matrices to compare three quantitative measures of spasticity (RTA, Hdf/Hctrl, and Hvib/Hctrl) with the MAS to determine whether, individually or collectively, they might be a reliable objective measure of spasticity, consistent with the standard clinical assessment and equally sensitive to cryotherapy-induced changes. They found that the RTA moderately correlated with the MAS and, when combined with the other quantitative measures into a linear composite variable, accounted for 38% of MAS score variance. In this study, we have used the existing dataset from Allison and Abraham’s study to calculate resistance torque (RT) scores, correlating them with the MAS scores and other quantitative measures. By obtaining the correlation coefficients, we hope to better understand the relevance and validity of RT as a clinically useful mechanical measurement of plantarflexor spasticity.

**METHODS**

**Research Design**

This was a correlational study. A quantitative biomechanical measure of spasticity, RT, was computed from an existing dataset and was correlated with three quantitative measurements and one clinical measurement (MAS) that had been reported previously.\(^11\) Correlation coefficients were computed before and after a 20-min application of ice packs to the calf muscles.

**Subjects**

Data collected earlier from 34 adult patients with traumatic brain injury with varied degrees of plantarflexor spasticity was used in this study. The subjects were selected from a pool of volunteers at the Healthcare Rehabilitation Center in Austin, TX. From a total of 60 ankles measured, data from 59 ankles were used in this study because of incomplete data from one ankle.
Data and Instrumentation

The MAS, RTA, soleus Hdf/Hctrl, and soleus Hvib/Hctrl data were collected as described previously. The torque data were collected from the patients by a torque motor apparatus, which also imposed the muscle stretch. This was a magnetic servomotor (Parker Compumotor, Inc. model 1030B) digitally controlled through a motor controller (model JSI) and driver (Dynaserve drive), which responded to commands relayed from a manually controlled Macintosh LC II desktop computer. Ramp and hold perturbations were provided by the torque motor, which also provided analog output signals corresponding to torque. On command from the desktop computer, the torque motor provided a ramp and hold stretch of the triceps surae muscles by rotating the foot at the ankle from 40 degrees plantarflexion to 0 degrees neutral position. A rotational acceleration of 300 revolutions per squared second and an average angular velocity of 235 degrees per second were maintained for most of the movement. This procedure was repeated 10 times for each patient before and after a 20-min cold pack application to the calf region. A 20-sec rest period was given between repetitions.

The computer control system was configured to simultaneously activate the torque motor and begin data collection. The analog output signals (EMG, torque, position) were converted to digital signals by an analog-to-digital converter board (LabLC) and then were collected in the Macintosh LC II computer. The computer was equipped with controlling software (LabVIEW version 2.2, National Instruments) with a sampling rate of 2 KHz, and the data were stored for offline analysis.

Data Analyses

The output torque was displayed in a graphical form. The ramp and hold perturbations provided by the torque motor evoked reflex responses that altered the otherwise smooth rise of the torque graph between the starting and the ending (accelerating and decelerating) pulses of the perturbation. The RT score was obtained by calculating the time integral of the torque curve between the two torque pulses. The ten measurements of the RT (one from each trial) thus obtained were averaged to provide one representative RT score for each ankle (Fig. 1).

Torque records were analyzed in this study to generate a mathematical measure that would satisfactorily measure the cumulative resistance. The torque measure was then correlated with the MAS scores of the patients in trials before and after cryotherapy (pre- and post-ice trials). The torque measure also was correlated with the quantitative measures generated earlier in the study. Two correlation matrices were created: one for pre-ice trials and the other for post-ice trials.

Statistical Analyses

Descriptive statistics were calculated for the RT variable. Correlation and regression calculations were performed with EPISTAT, version 5.0. Alpha levels were set to 0.1. For multiple regressions, full stepwise procedures were used; percent confidence intervals were set at 95%. Multiple regression was used to assess the ability of the RT variable to predict MAS scores in both pre- and post-ice trials. This was performed by adding the RT variable to the regression equation computed earlier using the other three quantitative vari-
Specifically, it was determined whether adding the RT variable to the multiple $R$ calculated earlier would improve the predictive accuracy of the regression model. Also, it was determined whether substitution of RT for RTA in the regression equation would alter the multiple $R$ in any significant manner.

**RESULTS**

Descriptive statistics for the RT variable are presented in Table 2. The range of MAS scores obtained in this study tended to be clustered toward the lower end. There were more ankles with lower scores of MAS and fewer ankles with higher MAS scores. This result has been described in further detail by Allison and Abraham. All measures of spasticity including MAS scores, RT, and RTA scores tended to improve with cryotherapy. The correlation matrices of Pearson coefficients between individual pairs of measures for the pre- and post-ice trials are presented in Tables 3 and 4. The coefficient for correlation between MAS score and RT score was 0.41 for the pre-ice trials and 0.42 for the post-ice trials. The Pearson coefficient for correlation between RTA score and RT score was 0.66 ($P < 0.001$) for the pre-ice trials and 0.75 ($P < 0.001$) for the post-ice trials.

The multiple $R$ value for the correlation between MAS scores and the linear combination of RTA, Hdf/Hctrl, Hvib/Hctrl, and the RT scores was 0.62 for the pre-ice trials. RT, the last included variable in the equation, was included at a $P$ value of 0.20. The multiple $R$ value for the correlation between MAS scores and the linear combination of RTA, Hdf/Hctrl, Hvib/Hctrl, and RT scores was 0.62 for the post-ice trials. RT, the last included variable in the equation, was included at a $P$ value of 0.27.

The study also was intended to determine whether inclusion of the RT variable in the regression equation used in the previous study (using the three identified quantitative variables to test their ability to predict the MAS) would increase the predictability or, in statistical terms, the multiple $R$. The multiple $R$ did not increase significantly with the inclusion of the RT variable in the regression equation. The linear composite variable thus obtained accounted for 38.50% of MAS score variation in the pre-ice trials and for 38.22% variation in the post-ice trials.

The inclusion of the RT variable in the regression equation was at nonsignificant $P$ values of 0.20 and 0.27 in the pre- and post-ice trials, respectively. This probably was primarily attributable to the phenomenon of multicollinearity exhibited between RT and the RTA variables. The moderate correlation between these two variables in both sets of trials suggests that the two variables might measure similar components of spasticity. This was further tested by replacing RTA with RT in the original regression equation. The multiple $R$ thus obtained for the pre- or post-ice trials did not change significantly.

The RT variable was included in the equation within 95% confidence limits, and the multiple $R$ obtained by the linear combination of RT, Hdf/Hctrl, and Hvib/Hctrl scores was 0.61 in the pre-ice trials and 0.59 in the post-ice trials. These $R$ values were comparable with the $R$ values obtained with RTA in the regression equation instead of RT. These results suggest that RT and RTA might measure similar mechanisms of spasticity.

**DISCUSSION**

The results show that the correlation coefficients between RT scores and MAS scores in both

| TABLE 2 Descriptive statistics for resistance torque (RT) variable |
|-------------------|-------------------|
|                   | RT Before Ice     | RT After Ice      |
| Mean              | 0.1556            | 0.154             |
| Standard deviation| 0.0371            | 0.0439            |
| Range             | 0.14939           | 0.20863           |
| Minimum           | 0.09417           | 0.08792           |
| Maximum           | 0.24356           | 0.29655           |

| TABLE 3 Correlation matrix for pre-ice trials |
|-------------------|-------------------|-------------------|
|                   | Hdf/ Hctrl       | Hvib/ Hctrl       | MAS    | RTA    |
| Hdf/ Hctrl        | 0.2959            | 0.1334            | 0.4184 | 0.655  |
| Hvib/ Hctrl       | 0.1334            | 0.489             | 0.4576 | 0.2273 |
| MAS               | 0.4184            | 0.4576            | 0.4508 | 0.4213 |
| RTA               | 0.655             | 0.2273            | 0.2075 | 0.4213 |

| TABLE 4 Correlation matrix for post-ice trials |
|-------------------|-------------------|-------------------|
|                   | Hdf/ Hctrl       | Hvib/ Hctrl       | MAS    | RTA    |
| Hdf/ Hctrl        | 0.068             | −0.027            | 0.5729 |        |
| Hvib/ Hctrl       | −0.027            | 0.4099            | 0.343  |        |
| MAS               | 0.4155            | 0.4099            | 0.343  |        |
| RTA               | 0.7524            | 0.1906            | 0.0425 | 0.4852 |

RT, resistance torque; Hdf/Hctrl, H dorsiflexion/H control amplitude ratio; Hvib/Hctrl, H vibration/H control amplitude ratio; MAS, modified Ashworth scale; RTA, reflex threshold angle. Boldface type, data generated by this study.
the pre- and post-ice trials were fair, suggesting a modest relationship between the two variables. The correlation coefficients between RT and RTA scores in both the pre- and post-ice trials were moderate, suggesting a very strong relationship between the two variables. The correlation coefficients between MAS scores and RT scores in both the pre- and post-ice trials were comparable, which might suggest that the treatment effect caused by cryotherapy had very similar effects on both variables.

RT scores can thus be used as an alternative to RTA as a measure of spasticity. The calculation of RTA requires sophisticated and reliable EMG measurement techniques. The process of arriving at an RTA score frequently involves visual inspection of the EMG trace to determine the onset of reflex activity, although some investigators have used objective criteria such as two or three standard deviations above baseline EMG amplitudes as an indicator of the onset of measurable activity. This process might lead to inaccuracies in measurement in the event of noisy EMG signals, and it requires a trained eye to detect the onset of reflex activity. The calculation of RT, on the other hand, is more straightforward, and the process could be readily automated to generate a numeric score in real time with each trial on a patient.

Large-amplitude ramp stretches of the ankle, such as those used in the current study, are similar to manually applied stretches used in the MAS and could be applied accurately and reliably using available equipment. The passive mode of an isokinetic dynamometer is already being used to measure spasticity in children with spastic cerebral palsy for research purposes by imposing controlled rotations of a joint while measuring the resistance to motion. However, torque measurements obtained in this study and from dynamometers in general do not differentiate dynamic tone from soft-tissue tightness, thus limiting their ability to evaluate spasticity per se.

The level of correlation required to validate the inclusion of RT as an accepted measure of spasticity is difficult to determine. The intrarater reliability of the MAS to assess plantarflexor spasticity has been shown to be 0.73, which barely meets minimal criteria for acceptable reliability. These studies of spasticity in the ankle plantarflexors indicate that the percent agreement of MAS scores between testers is only 55% for patients with traumatic brain injury and approximately 40–50% for patients with spinal cord injury. Taking this into consideration, the Pearson coefficient in our study of 0.42 between RT and MAS scores can arguably be interpreted more favorably than the correlation coefficients might suggest at the outset.

RT was calculated as the time integral of the torque graph between the starting and ending torque pulses of the ramp and hold stretch. RT would thus be a measure of the cumulative history of the torque during the imposed stretch. The MAS is a qualitative, subjective method of measuring the cumulative resistance perceived by the clinician during the imposed passive stretch. Thus, it could be argued that RT is a mathematical, quantitative method of assessing spasticity in a similar way that the MAS would assess it qualitatively, despite the limitations of both techniques in their inability to differentiate between static and dynamic components of resistance. Also, RT can be used for augmenting the MAS assessment of spasticity.

Katz et al. correlated MAS scores with RT and torque at a fixed angle during ramp and hold stretches for the elbow flexors. They obtained significant high correlation between RTA and MAS scores, but they did not obtain a significant correlation between torque and MAS score. This angle was chosen arbitrarily, and it was acknowledged in the study that a “different torque measure might be more useful as a measure of relative change in tone since intrinsic mechanical properties of the limb should stay constant over time.” Katz and Rymer have expressed concern about the variability of torque in different limbs in the same subject and in different subjects. More than the absolute measurement of torque, the question that needs to be answered is whether the measured torque changes reliably with changes in spasticity. This study suggests that because RT had a similar correlation with the MAS after treatment with cryotherapy, it is arguably a good measure of spasticity.

The qualitative, subjective descriptions of muscle hypertonia given by the clinical scales necessarily depend on the examiner’s perception. Currently, physicians usually have very little difficulty in the diagnosis of spasticity in most of their patients, but the problem arises when quantitative consideration must be added, probably because spasticity is not a simple entity but a syndrome originating from a variety of disorders that express disturbances over spinal motor neurons. The MAS remains in common use, secondary to ease of use, despite its obvious functional limitations. With an increasingly greater range of treatment options available for spasticity and an increase in the number of centers participating in studies of interventions for spasticity, published reports reflect increasing use of objective quantification techniques. Reliable quantitative assessment also facilitates communication between clinicians and is useful in documenting clinical rehabilitation progress.

Mechanical measures of spasticity, such as the RT obtained in this study, can increase the objectivity, repeatability, and precision of spasticity quantification and can provide additional insight
into neuropathology. The relationships between the available qualitative scales and quantitative measures have to be understood more thoroughly for the latter to attain acceptability in clinical practice. Such a potentially significant relationship has been demonstrated in this study.

**Limitations of the Study**

Although the correlation between RT and MAS scores was significant and favorably directed to suggest that it could be an appropriate indicator of spasticity, the overall magnitude of the correlation was not sufficient to support its immediate acceptance in clinical practice. Squaring the correlation coefficients from Tables 2 and 3 indicates that about 18% of the variance in MAS scores could be accounted for by the RT variable.

The fair level of correlation in this study could be attributable to the lower reliability of MAS at the ankle, as shown in the previous study, and because of the restricted range of MAS scores in the subjects used. Stronger relationships would likely have been demonstrated if patients with a larger range of spasticity had been tested and if more subjects with higher grades of spasticity, who could have demonstrated a better treatment effect with cryotherapy, had been included in the study. Further, this study only employed one velocity of stretch, which may have limited the findings.

The use of torque measurements as an index of spastic hypertonia is certainly not a straightforward matter, especially because most clinicians are inexperienced with mechanical measurements of any kind. Moreover, the absolute torque recorded in the limbs of different subjects will certainly vary with limb mass, muscle bulk, and the characteristics of each individual’s muscular anatomy. This implies that objective quantification of hypertonia in a diverse group of spastic subjects is unlikely to be completely successful—not, at least, until we document the range of torque variation in normal passive limbs.

**CONCLUSIONS**

RT is a measure of the cumulative history of the torque during an imposed stretch of a joint in a spastic extremity and is a mathematical, quantitative correlate of the clinical scale, MAS, in assessing plantarflexor spasticity. The reliability and validity of the RT as a measure of spasticity needs to be tested in future studies. Although RT and other mechanical measures of spasticity offer many advantages, much research needs to be done to establish a standard measurement of spasticity that can be implemented easily in the clinic.

**REFERENCES**

21. Allison SC, Abraham LD: Sensitivity of qualitative and...
quantitative spasticity measures to clinical treatment with cryotherapy. *Int J Rehabil Res* 2001;24:15–24


Motor Neuron Inhibition–Based Gene Therapy for Spasticity

ABSTRACT


Spasticity is a condition resulting from excess motor neuron excitation, leading to involuntary muscle contraction in response to increased velocity of movement, for which there is currently no cure. Existing symptomatic therapies face a variety of limitations. The extent of relief that can be delivered by ablative techniques such as rhizotomy is limited by the potential for sensory denervation. Pharmacological approaches, including intrathecal baclofen, can be undermined by tolerance. One potential new approach to the treatment of spasticity is the control of neuromuscular overactivity through the delivery of genes capable of inducing synaptic inhibition. A variety of experiments in cell culture and animal models have demonstrated the ability of neural gene transfer to inhibit neuronal activity and suppress transmission. Similarly, enthusiasm for the application of gene therapy to neurodegenerative diseases of motor neurons has led to the development of a variety of strategies for motor neuron gene delivery. In this review, we discuss the limitations of existing spasticity therapies, the feasibility of motor neuron inhibition as a gene-based treatment for spasticity, potential inhibitory transgene candidates, strategies for control of transgene expression, and applicable motor neuron gene targeting strategies. Finally, we discuss future directions and the potential for gene-based motor neuron inhibition in therapeutic clinical trials to serve as an effective treatment modality for spasticity, either in conjunction with or as a replacement for presently available therapies.

Key Words: Spasticity, Transgene Expression, Motor Neuron, Gene Targeting

Spasticity is a condition resulting from excess motor neuron excitation caused by lesions in the upper motor neuron pathway that lead to the absence of inhibition of alpha and/or gamma motor neurons. This loss of inhibition results in involuntary muscle contraction, causing stiffness interfering with movement, speech, and locomotion.1,2 Afflicting more than 12 million people worldwide, spasticity is commonly caused by stroke, multiple sclerosis, cerebral palsy, cerebral infection (encephalitis/meningitis), and/or cerebrospinal trauma.3 The most common clinical manifestations of spasticity are hypertonicity, clonus, muscle spasm, and/or fixed joints.4–7 Symptoms can present either intermittently (triggered by pain or other stimuli) or continuously, and they occur most commonly in response to increased velocity of movement.8
MECHANISM OF SPASTICITY

The exact mechanism of spasticity in humans is incompletely understood, primarily because it is multifactorial in nature. It is generally understood that spasticity is caused by pathology involving the stretch reflex, which normally causes a muscle to contract to resist the force that is stretching it. For normal movement to occur, the brain must be able to selectively turn this reflex off, usually via inhibitory signals relayed to the spinal cord via the corticospinal tract. However, damage to this circuit results in disinhibition of the stretch reflex; over time, this reduces the triggering threshold until excessive and complete muscle contraction can occur even at rest, making the limb virtually impossible to move.

Specific causes proposed include (a) alpha motor neuron hyperexcitability resulting from an imbalance in excitatory vs. inhibitory alpha motor neurons, and (b) gamma motor neuron hyperactivity manifesting as increased sensitivity of muscle spindle to stretch (fusimotor hyperactivity). Additional causes involve damage to descending tracts that control interneurons responsible for (a) mediating presynaptic inhibition of the Ia terminals on the alpha motor neuron, (b) mediating type II afferents, and (c) reciprocal Ia inhibition. Such damage results in greater afferent stimulus to the alpha motor neuron as a result of stretch, decreased inhibition from type II afferents, and loss of normal inhibition of antagonist muscle during muscle stretch. Finally, a mechanism of decreased recurrent inhibition from Renshaw cells as a consequence of supraspinal damage has also been proposed. Whereas each individual cause contributes to the clinical picture observed in spasticity, it is unlikely that any single hypothesis is sufficient to explain the exact mechanism of spasticity.

EXISTING THERAPIES AND LIMITATIONS

Medical Therapies

Although there are a number of oral medications available to treat spasticity, almost none are effective without significant side effects. The most common medications are diazepam, baclofen, and progabide, all of which are designed to increase presynaptic inhibition of alpha motor neurons by activation of γ-aminobutyric acid (GABA) receptors. However, these treatments are associated with a high incidence of adverse effects, including sedation, weakness, fever without infection, and elevated liver enzymes. Agents affecting ion flux in skeletal muscle, such as dantrolene, lamotrigine, and riluzole, share side effects of muscle weakness, sedation, and idiosyncratic hepatitis. Additional agents that act on monoamines, such as tizianidine, are associated with a similar side-effect profile. These significant toxicities limit the doses of medication that can be employed, thereby limiting efficacy. Furthermore, the issue of tolerance significantly hinders the long-term efficacy of any pharmacologic therapy, particularly for baclofen and diazepam.

An alternative medical therapy involves the use of clostridial toxin (i.e., botulinum toxin), which acts by decreasing acetylcholine release at the neuromuscular junction, resulting in a neuromuscular blocking effect. However, the results from this therapy are often transient, with redosing complicated by tachyphylaxis, and increasing dosage complicated by severe muscle weakness. The issue of cost is another consideration. The cost of clostridial toxin treatment might hinder its extensive clinical applications because conventional oral therapies are much less expensive.

Surgical Therapies

Surgery for spasticity is reserved for cases refractory to medical management or for those that cannot be medically managed because of intolerable side effects. The most common surgical options are generally orthopedic (consisting primarily of tendon-release operations) and neurosurgical. Neurosurgical procedures fall into two categories: nonablative and ablative. The most frequently used nonablative procedure is intrathecal baclofen (ITB), which is generally offered for refractory patients with chronic spasticity (>12 mos). To be considered a candidate, a patient must demonstrate a positive response to ITB at a test dose of less than 100 μg, compared with no response to placebo. Although highly effective in improving muscle tone and reducing postoperative spasticity, ITB is fraught with catheter- and wound-related morbidity, both at the time of implantation and throughout the life of the implanted device.

There are limited data characterizing the problem of tolerance to ITB. In many cases, ITB tolerance is ascribed to progression of the underlying disease (in amyotrophic lateral sclerosis and multiple sclerosis) or to dynamic catheter obstruction (kinking), which is difficult to demonstrate on standard pump contrast injections (pumpograms) or nuclear medicine studies. In our practice, we use inpatient externalized catheter ITB trials to address the question of baclofen tolerance. In this context, it is easy to assess the patency of the catheter and document the threshold for response to an intrathecal drug. Outpatient trials of bolus intrathecal injection can be attempted, but these are often misleading because of the inherent differences in the pharmacokinetics of bolus and pump injection. However, the majority of our patients require gradually escalated doses of ITB to maintain adequate control of spasticity.
Another nonablative option is spinal cord stimulation, which has been shown to facilitate spasticity control, spasm inhibition, and gait improvement in spastic patients,27–29 likely by selective modification of segmental spinal reflexes.30 However, efficacy varies greatly, is highly dependent on both electrode position and the degree of stimulation, and has questionable cost-effectiveness.31,32 Furthermore, spinal cord stimulation is fraught with device-related morbidity related to infection, electrode migration, wire breakage at the connector site, skin breakdown over the lead, and development of high impedance.33

The most common ablative procedure is selective dorsal rhizotomy, which uses intraoperative electromyography and stimulation to identify the rootlets most responsible for causing severe spasticity.34,35 Selective dorsal rhizotomy has consistently been proven to reduce spasticity of cerebral palsy in children, with earlier age at surgery associated with a reduced incidence of lower-extremity deformities (i.e., contractures, secondary skeletal torsion) requiring orthopedic surgery later in life.36–40 However, selective dorsal rhizotomy comes with morbidity as well, related both to the risk of surgery and the reported increased incidence of hip subluxation after treatment.41 Furthermore, selective dorsal rhizotomy has not been shown to impact alpha motor neuron–induced spasticity.34,35 Less common ablative procedures include percutaneous radiofrequency rhizotomy and surgical myelotomy (dorsal root entry zone procedures), each of which provides moderate efficacy in exchange for a permanent central nervous system (CNS) lesion and possible supplementation with subsequent ablative procedures.42–45

POSSIBLE FUTURE THERAPIES

Despite the wide range of medical and surgical treatments for spasticity, there is currently no treatment modality that is hardware free, reversible, adjustable, nondestructive, and not subject to tolerance.2,3,14–16,46 A potential modality for satisfying these criteria is the use of viral vectors to elicit effects on muscle contraction via gene transfer, because selective control of certain genes has been shown to modulate neuronal activity in multiple applications.47–52 Typically, motor neurons communicate with muscle fibers by releasing acetylcholine at the neuromuscular junction. The release of acetylcholine causes an excitatory postsynaptic potential in the muscle fiber that triggers a postsynaptic action potential, which then causes the muscle fiber to contract. Because spasticity results from excess excitation of motor neurons, transgene-induced inhibition of motor neuron excitatory responses could alleviate or even abolish the clinical manifestations of spasticity while evading the problems associated with tolerance. It is critical to achieve motor neuron–specific gene transfer to achieve targeted therapeutic effects in spasticity treatment. With the development of vector targeting technology, a specific cell population, such as motor neurons, can be transduced selectively. This issue will be discussed in detail below.

POTENTIAL INHIBITORY TRANSGENES

One of the most widely studied inhibitory transgenes is the gene encoding glutamate decarboxylase (GAD), the rate-limiting enzyme required to produce the inhibitory neurotransmitter GABA. In vitro and in vivo studies using retroviral vectors and adeno-associated virus (AAV) vectors have suggested that it is feasible to achieve long-term GAD expression in the CNS.53–56 Studies also have shown that GAD expression in the CNS induces GABA production in vector-transduced cells. Transfer of GAD using viral vectors therefore bears the potential application in disorders resulting from overexcitation in the nervous system. In a rat Parkinson disease model, researchers have transferred GAD genes in an AAV vector into the glutamatergic neurons of the subthalamic nucleus.57 GAD was expressed, and the expression of GAD induced production of GABA in these neurons. These findings suggest that AAV-mediated GAD gene transfer might provide a treatment option for overactive diseases such as Parkinson disease.57,58 These results have led to an ongoing Phase I trial of GAD gene transfer to the human subthalamic nucleus for medically refractory Parkinson disease patients.58 In addition to GAD, two other transgenes have emerged as candidates that are potentially capable of providing the inhibitory impact on motor neurons necessary for treatment of spasticity. The performance of these transgenes in previous work has generated optimism for their potential in combating spasticity.

Tetanus Toxin Light Chain

Our laboratory has focused on the gene for the light chain (LC) fragment of clostridial neurotoxin. The expression of this gene in neurons provides inhibition of synaptic function in transgenic mice via reversible suppression of glutamatergic neurotransmission.59,60 Clostridial intoxication in neurons involves the production of inactive single-chain clostridial neurotoxin polypeptides, which are released after bacterial lysis. This release converts the polypeptide from an inactive single-chain molecule to an active di-chain molecule composed of a heavy chain (HC) and an LC fragment linked by a single disulfide bond. HC binds axon terminals and triggers internalization of the toxin. Once inside the neuron, reduction of the disulfide bond releases the active LC fragment. The activated LC cleaves the soluble N-ethylmaleimide–sensitive factor attachment receptor proteins responsible for synaptic vesicle membrane fusion.61–63 Reduction in vesicle fusion inhibits neurotransmission without inducing neuronal cell death.59,61,62 Recent in vitro
work in our laboratory has demonstrated that biologically active clostridial LC proteins can be successfully produced in cultured cells through viral gene transfer. The in vitro expressed LC protein was able to digest synaptobrevin, a soluble N-ethylmaleimide-sensitive factor attachment receptor protein that is involved in neurotransmitter release. Our in vivo experiments demonstrated that injection of an LC-expressing adenoviral vector into the rat spinal cord or the dSC nucleus in the brain stem inhibited neurotransmitter release. Our in vivo experiments demonstrated that injection of an LC-expressing adenoviral vector into the rat spinal cord or the dSC nucleus in the brain stem inhibited neurotransmitter release. The surrounding CNS structures were not found to be affected by LC gene expression during the 1-mo observation period.59 Such specific, reversible, effective neuronal inhibition makes LC a viable transgene candidate for exploring gene-based treatment of spasticity.

Inwardly Rectifying Potassium Channel Kir2.1

Another gene of interest as a feasible modulator of motor neuron activity is Kir2.1, which encodes inwardly rectifying potassium (Kir) channels in the heart and brain. For inwardly rectifying potassium channels, the inward flow of potassium ions at subthreshold is greater than the outward flow of potassium ions for the opposite driving force. This inward rectification results when intracellular magnesium ions and polyamines enter the ion channel pore from the cytoplasmic side but are unable to pass through it to the extracellular solution. The block is more intense at decreased membrane potentials as the larger depolarization facilitates the movement of magnesium ions and polyamines into the pore. In contrast, as membrane potentials approach the resting membrane potential, the decreasing depolarization hinders magnesium ion movement into the pore. When the membrane potential exceeds the resting membrane potential (hyperpolarization), magnesium ions become prevented from entering the channel. Inwardly rectifying potassium channels prevent the membrane potential from depolarizing by increasing the membrane potassium conductance. This increase in potassium permeability counterbalances the excitatory synaptic potentials that drive the initial membrane depolarization, hence inhibiting the formation of the action potential. In this way, the Kir2.1 contributes to stabilizing the resting potential at a sufficiently negative level to prevent enough sodium channel availability for action potential in the CNS and heart. Kir2.1 has been demonstrated to inhibit both evoked and spontaneous activity of neurons in vitro.64

Thus, overexpression of inwardly rectifying potassium channels in motor neurons has the potential to inhibit depolarization, resulting in a subsequent drop in action potential generation and inhibition of neuromuscular transmission (Fig. 2). Previous work has demonstrated that Kir2.1 can be successfully transferred into cultured neurons, resulting in selective, inducible, reversible genetic expression of neuronal excitability.64 For these reasons, the gene for Kir2.1 and other inwardly rectifying channels may provide a means for the control of motor neuron overactivity, thus providing an approach to spasticity.

STRATEGIES FOR CONTROL OF TRANSGENE EXPRESSION

Inducible Gene Expression System

Gene therapy will only prove beneficial as a treatment modality for spasticity if it provides advantages over existing pharmacologic and lesion-based modalities. For this to be accomplished, transgene expression must be both adjustable and reversible. One way to control viral vector-mediated transgene expression is to use inducible promoter elements. Several inducible promoter systems have been developed for this purpose, such as tetracycline, RU-486, rapamycin responsive systems, and the chimeric drosophila/bombyx ecdysone receptor system.65–68 As an illustration of how
inducible systems work, we will discuss the tetracycline responsive promoter system, of which there are two types. The first can be turned on in the presence of doxycycline (tet-on), and the second can be turned off in the presence of doxycycline (tet-off). In the tet-on system, the transgene is under the control of cytomegalovirus immediate early (CMVie) promoter, which is composed of a cytomegalovirus (CMV) minimal promoter and seven repeat sequence from a bacterial tet repressor DNA binding sequence (tetO). The other component of the tet-on system is a cassette that expresses tetracycline responsive transactivator proteins. In the presence of doxycycline, the complex formed between tetracycline responsive transactivator protein and doxycycline binds to CMVie promoter to turn on downstream therapeutic transgene expression.

As alluded to above (and true for other inducible systems), two components are required to achieve regulated gene expression. Because some of the compounds used to induce gene expression for these systems can pass the blood–brain barrier, they can be applied to the control of gene expression in the CNS. As such, they hold promise as a means to control transgenes expressed in the spinal cord broadly and motor neurons specifically, rendering them viable for the treatment of spasticity. Systemic delivery and intrathecal infusion of these trigger compounds have been proposed as means to control antispasticity transgenes. Because the candidate genes to treat spasticity encode proteins that either shut down neurotransmission (LC) or induce hyperpolarization (Kir), an inducible system could potentially limit undesirable side effects of gene expression. Furthermore, the use of an inducible expression system may

FIGURE 2 Mechanism of action of Kir2.1 in motor neurons. A, Typical motor neuron with voltage at resting potential. The neurotransmitter binds, resulting in an influx of cations (sodium and calcium), driving the voltage toward the threshold as depolarization occurs. B, The excitatory postsynaptic potential reaches its threshold and induces opening of voltage-gated channels, resulting in a large influx of sodium ions, triggering the action potential. C, Motor neuron at resting potential with Kir2.1 introduced. Kir2.1 overexpression clamps the resting membrane potential to the reversal potential of K+ ions and makes it more resistant (relative to A) to depolarizing forces caused by neurotransmitter binding. D, The excitatory postsynaptic potential displaces the magnesium ion (Mg2+) as Kir drives repolarization. The Kir-driven eflux of K+ balances the influx of Na+, thereby preventing the voltage from reaching its threshold, and keeping the voltage-gated channels closed. This inhibition allows for control of neuromuscular transmission.
provide a way to adjust gene expression levels according to symptom severity.

**MOTOR NEURON TARGETING STRATEGIES**

A variety of strategies exist for targeting gene delivery to specific cell populations. Vectors can be designed to specifically bind to neurons in general or to motor neurons specifically. Further, cell type–specific promoters can be used to restrict gene expression to defined neuronal populations. As discussed earlier, existing pharmacologic therapies for spasticity (including ITB) have been limited by off-target effects that create side effects at higher doses. These off-target effects result from the binding of these agents to neurons in a variety of functional systems other than the desired motor neuron and spasticity-inducing reflex. The ideal scenario, therefore, is the use of the most minimally invasive approach that would permit the introduction of an efficient, controllable gene into the neurons controlling the spasticity-inducing reflex. Spinal cord motor neurons are the primary targets for gene delivery in the treatment of spasticity. Both promoter-level motor neuron targeting, and enhanced motor neuron binding and uptake through vector capsid modifications, are currently being pursued.

**Direct Injection**

Animal experiments have demonstrated that direct spinal cord injection of viral vectors is a feasible, albeit risky, way to introduce genetic material into the spinal cord. Furthermore, our laboratory and others have reported gene transfer to motor neurons via direct injection of a wide variety of viral vectors into the animal spinal cord. However, diffuse expression of transgenes that cause neural inhibition would be predicted to affect sensory systems as well as motor systems. Such a diffuse effect would be acceptable in the context of the treatment of spastic paraplegia, but not for patients with preserved sensory and motor function. Another concern related to direct spinal cord injection is the potential disinhibition of motor neurons attributable to inhibitory gene expression in spinal cord interneurons. As such, effective motor neuron targeting is critical for the treatment of spasticity using gene transfer. Currently, two strategies exist to achieve motor neuron–specific gene transfer. The first strategy is to genetically modify viral vector surface proteins with a short motor neuron–specific peptide to render these vectors motor neuron specific. The second strategy is the application of motor neuron–specific promoters to confine gene expression to motor neurons.

A variety of tissue-specific and cell-specific promoters have been tested for gene transfer. Myelin basic protein is expressed in oligodendritic cells and Schwann cells in vivo. Previous studies have shown that myelin basic protein–directed oligodendritic–specific green fluorescent protein gene expression occurs mainly in the white matter but not in other cell types such as neurons, astrocytes, or microglial cells. Other tissue-specific promoters including neuro-specific enolase promoter, platelet-derived growth factor β-chain promoter, or β-glucuronidase promoter have also been studied for tissue-specific gene expression.

Although similar promoter-level restriction of gene expression has yet to be achieved for motor neurons, the study of cell differentiation has led to the recognition of several genes that are specifically expressed in neural progenitors as they mature into motor neurons. The identification of the promoters that control the expression of these genes provides a tool to achieve motor neuron–specific therapeutic gene expression. HB9, a homeodomain transcription factor, is expressed selectively by postmitotic spinal motor neurons in the developing vertebrate CNS and serves as a marker for the motor neuron phenotype. A 9-kb HB9 promoter sequence from the 5′ Hb9 gene has been shown to direct motor neuron–specific gene expression in vivo studies. Furthermore, Nakano et al. have isolated a 125-bp enhancer sequence from the homeobox gene promoter (HB9) region that restricted gene expression to the spinal cord motor neurons in transgenic animals. This enhancer/promoter sequence will be particularly useful because it is small enough to be accommodated in almost all available gene-transfer vectors. Interestingly, Pramatarova et al. have reported that neurofilament LC promoter directed spinal cord motor neuron–specific gene expression of superoxide dismutase 1 mutation G37R. We anticipate that this same promoter could be used to drive motor neuron–specific gene expression for the treatment of spasticity in the future. Finally, the promoters for the genes that control acetyl choline production and metabolism can be leveraged to design motor neuron–specific expression systems.

Although promising because of its ability to achieve high levels of gene expression, the delivery method of direct injection carries with it two major problems. The first is that it requires stable gene expression for a long period of time because repeated invasive surgery is not desirable. Secondly, the possibility of spinal cord trauma from direct spinal cord injection represents a serious source of morbidity. For these reasons, design of vectors capable of enhanced retrograde axonal transport (remote delivery) has been investigated as an alternative to the direct injection strategy.
Remote Injection

Remote delivery of therapeutic vectors can bypass direct CNS trauma by vector injection into innervated muscle groups and peripheral nerves. This delivery method has the advantage of facilitating repeated application with relatively low risk to the patient. Once injected, the foreign genetic material can be ferried into lower motor neurons of the spinal cord via retrograde axonal transport.\(^82,80\)

Skeletal muscles are innervated by nerve fibers from lower motor neurons in the spinal cord or brain stem. Active material transport between the nerve terminals and the cell body in both anterograde and retrograde directions exists to support the metabolic needs of these remote terminals. Therefore, it is theoretically possible to inject appropriately designed therapeutic agents into the muscle and have the injected agents transported back to the cell body through retrograde axonal transport. A variety of viruses and toxins have evolved to capitalize on this conduit into the CNS. For example, the rabies virus can be transported to CNS neurons after peripheral inoculation, and the herpes virus can be transported into ganglion neurons from peripheral inoculation sites.\(^81\) Clostridial tetanus toxin also can be retrogradely transported to the CNS neurons from peripheral inoculation sites. Not surprisingly, these properties all have been tested in gene-transfer vector design to achieve central targeting from peripheral inoculation.

Axon terminal uptake of the rabies virus at the neuromuscular junction depends on the rabies G glycoprotein.\(^82\) Consequently, efforts have been made to pseudotype or coat gene-transfer vectors with this molecule. Mitrophanous et al.\(^83\) have reported the development of an equine infectious anemia virus–based lentiviral vector pseudotyped with rabies G proteins for neuron-specific transduction. These vectors have been demonstrated to undergo avid uptake in innervated muscle fibers and enhanced retrograde transport into the related motor neurons in a retrograde fashion.

As mentioned above, tetanus toxin is comprised of an HC and an LC and undergoes retrograde axonal transport to CNS neurons from peripheral inoculation. The HC component binds to its receptor GT1b and mediates the retrograde migration of the holotoxin. This property has been used to ferry therapeutic or tracer agents to CNS neurons, as the HC component has been used to deliver therapeutic agents such as superoxide dismutase 1 and cardiotropin-1 into motor neurons in attempts to target motor neuron diseases.\(^84–87\)

Despite these promising results, with currently available vector systems, only a low degree of gene-transfer efficiency can be achieved. Because a large amount of vectors are needed to generate a favorable clinical benefit, this delivery method may increase the antigenic burden to the patients in addition to posing a manufacturing challenge. Additionally, because retrograde transport impacts the dorsal root ganglion as well as the desired motor neuron, potentially unwanted side effects may occur, as with the current modality of clostridial toxin therapy previously discussed.

With these limitations in mind, a discussion of gene therapy remains important for the treatment of spasticity because (a) gene delivery can control synaptic function, (b) this control can be specifically targeted to motor neurons, and (c) the expression of these genes can be controlled once applied to motor neurons. The development of a modality that is rapidly progressing to conquer these frontiers will, in the near future, mark a dramatic advance in the care that physicians can provide to patients with spasticity.

Viral Gene-Transfer Vectors

As alluded to above, gene-transfer vectors currently used in preclinical and clinical studies are mainly viral vectors. It is therefore appropriate to describe briefly the characteristics of some major viral vectors. Retroviral vectors, one of the earliest types of viral vectors developed for gene transfer, are still widely used in preclinical and clinical studies. Recombinant retroviral vectors can generally accommodate up to 8 kb of transgene. This group of viral vectors only infects dividing cells. Theoretically, they mediate long-term gene expression as the transgene is integrated into the host-cell genome. For the same reason, retroviral vectors may be tumorigenic because the transgene sequence and the viral long terminal repeat sequences may integrate randomly and activate a tumor-suppressor gene.\(^88\) Adenoviral vectors are another group of vectors that are being used for gene transfer. Conventional adenoviral vectors can take about 7 kb of transgene sequence. The newer, gutless adenoviral vectors can accommodate up to 35 kb of foreign DNA sequence. Gene expression from adenoviral vectors is usually transient. The disadvantage of adenoviral vectors is that these vectors have the potential to trigger deadly immune responses.\(^89\) AAV also can be genetically modified as a gene-transfer vector. AAV vectors can hold up to 5 kb of foreign DNA sequence. Gene expression from AAV vectors is usually long-lasting relative to adenoviral vectors. Although wild-type AAV integrates into a specific site on chromosome 19, recombinant AAV might lose this ability; therefore, it is not considered tumorigenic. The downside of AAV vectors is their relatively small packaging capacity of 5 kb, as mentioned above. Newer techniques, however, exist to increase their packaging capacity to more than 5 kb.\(^90\) So far, AAV vectors have a good safety
FUTURE DIRECTIONS

All of the transgenes discussed in the present review inhibit neuronal and synaptic activity in a relatively nonspecific fashion. As the mechanisms of spasticity are further elucidated, we anticipate that new candidate transgenes will emerge that will be capable of specifically blocking spasticity as opposed to general neuronal activity. The continued evolution of gene-transfer technology has provided the ability to transfer desired genetic material into human neurons. Practical motor neuron gene therapy for spasticity will depend on means for safe, durable, controllable, and specific gene delivery. Although an acceptable system that incorporates all of these features is not yet available, in the present review we have attempted to demonstrate progress in each of these dimensions. The continued dire need for motor neuron protection strategies for the treatment of amyotrophic lateral sclerosis and spinal muscular atrophy is driving the evolution of techniques for safe, durable, and specific motor neuron gene delivery. Similarly, the evolution of systems for controlled gene expression is being driven by the need for this feature in the therapy of a wide variety of diseases. The application of gene-based neuronal inhibition to the control of subthalamic nucleus overactivity in Parkinsonian humans presages the broader application of this approach to epilepsy, spasticity, pain, and other functional disorders of the nervous system. We anticipate that physicians will play an intimate role in the deployment of these therapies.

ACKNOWLEDGMENTS

We would like to thank Dr. Thais Federici for invaluable assistance.

REFERENCES


50. Nakano T, Windrem M, Zappavigna V, Goldman SA: Iden-


Spinal Injury Mimicking Lumbosacral Plexopathy

Jennie-Corinne Baublitz-Brenenborg, DO, Michael J. DePalma, MD, and David X. Cifu, MD

From the Department of Physical Medicine and Rehabilitation, Virginia Commonwealth University/Medical College of Virginia, Richmond, Virginia.

A 51-yr-old female presented with a chief complaint of left lower-limb pain and weakness after falling from a ladder. She experienced immediate lumbar pain that resolved during the ensuing week. One month later, she experienced an acute onset of stabbing left anterolateral thigh, anterior leg, and plantar foot pain, with dorsal foot paraesthesias. Physical examination revealed diminished sensation over the left first dorsal webspace, 4+/5 left ankle dorsiflexion, knee extension, hip flexion, hip abduction, and 4/5 great toe extension, and a diminished left medial hamstring reflex. Left straight-leg raise to 45 degrees and reverse straight-leg raise to 90 degrees of knee flexion reproduced radicular pain.

Electrodiagnostic evaluation revealed +1 positive sharp waves and fibrillation potentials in the left tensor fascia lata and vastus medialis, +1 positive sharp waves in the tibialis anterior, and peroneus longus with 15-Hz recruitment frequencies of increased duration, normal amplitude motor unit potentials. No abnormalities were noted in distal foot musculature, corresponding right lower-limb musculature, or lumbar paraspinals. The left saphenous sensory nerve action potential amplitude was reduced by approximately 75%, and the left extensor digitorum brevis compound muscle action potential amplitude was 50% attenuated compared with the right. Contralateral nerve-conduction studies were normal.

The electromyogram is an indirect functional representation of the axon, reflecting altered muscle cell membrane physiology.1,2 The sensory nerve action potential can localize the lesion either proximally to the dorsal root ganglion (DRG), such as in lumbosacral radiculopathy, with preserved sensory nerve action potentials, or distally to the DRG, with diminished or absent sensory nerve action potential. In our case, a lesion at or distal to the DRG was suspected on the basis of the electrodiagnostic data. A pelvic magnetic resonance imaging (MRI) did not reveal a structural origin for lumbar plexopathy. Subsequent contrast-enhanced lumbosacral MRI demonstrated a lateral extrusion of the L4–5 disc effacing the left L4 DRG and displacing the left L5 nerve root (Fig. 1). This structural abnormality corroborated both the clinical presentation and the electrophysiologic evidence.

MRI is useful in defining nerve root morphology because it can differentiate between nerve and surrounding soft tissue. However, the morphologic features of DRGs are difficult to evaluate on MRI because of variable position and small size.3 Therefore, MRI studies must be evaluated attentively in a patient with clinical evidence of disc herniation, particularly when DRG involvement is suspected or suggested by electrodiagnostic evaluation.4

REFERENCES


All correspondence and requests for reprints should be addressed to Michael J. DePalma, MD, Department of Physical Medicine and Rehabilitation, Virginia Commonwealth University/Medical College of Virginia, 1222 Whiby Road, Richmond, VA 23227.