1. Masthead • MISCELLANEOUS  
   Page A2

2. Editorial Board • EDITORIAL BOARD  
   Page A3

3. Table of Contents • CONTENTS LIST  
   Pages A9-A10

Original Articles

4. A Randomized Controlled Trial of Modified Constraint-Induced Movement Therapy for Elderly Stroke Survivors: Changes in Motor Impairment, Daily Functioning, and Quality of Life • ARTICLE  
   Pages 273-278  
   Ching-yi Wu, Chia-ling Chen, Wen-chung Tsai, Keh-chung Lin and Shih-han Chou

5. Effect of Motorized Scooters on Physical Performance and Mobility: A Randomized Clinical Trial • ARTICLE  
   Pages 279-286  
   Helen Hoenig, Carl Pieper, Laurence G. Branch and Harvey Jay Cohen

6. Spasticity Experience Domains in Persons With Spinal Cord Injury • ARTICLE  
   Pages 287-294  
   Jane S. Mahoney, Joan C. Engebretson, Karon F. Cook, Karen A. Hart, Susan Robinson-Whelen and Arthur M. Sherwood

7. Motor Points for the Neuromuscular Blockade of the Subscapularis Muscle • ARTICLE  
   Pages 295-297  
   Tim P. Harrison, Anna Sadnicka and Deborah M. Eastwood
8. Poststroke Shoulder Pain: Its Relationship to Motor Impairment, Activity Limitation, and Quality of Life • ARTICLE
   Pages 298-301
   John Chae, Don Mascarenhas, David T. Yu, Andrew Kirsteins, Elie P. Elovic, Steven R. Flanagan, Richard L. Harvey, Richard D. Zorowitz and Zi-Ping Fang

   Pages 302-308
   Scott R. Millis, Don Straube, Cherdsak Iramaneerat, Jr, Everett V. Smith and Patrick Lyden

10. Clinimetric Properties of the Duruoz Hand Index in Patients With Stroke • ARTICLE
    Pages 309-314
    Nebahat Sezer, Gunes Yavuzer, Koncuy Sivrioglu, Pınar Basaran and B. Fusun Koseoglu

11. Aerobic Capacity After Traumatic Brain Injury: Comparison With a Nondisabled Cohort • ARTICLE
    Pages 315-320
    Kurt A. Mossberg, Danielle Ayala, Tracey Baker, Justin Heard and Brent Masel

12. The Development and Validity of the Salford Gait Tool: An Observation-Based Clinical Gait Assessment Tool • ARTICLE
    Pages 321-327
    Brigitte Toro, Christopher J. Nester and Pauline C. Farren

13. Inter- and Intraobserver Repeatability of the Salford Gait Tool: An Observation-Based Clinical Gait Assessment Tool • ARTICLE
    Pages 328-332
    Brigitte Toro, Christopher J. Nester and Pauline C. Farren

14. Estimated Prevalence of Obstructive Sleep Apnea–Hypopnea Syndrome After Cervical Cord Injury • ARTICLE
    Pages 333-337
    Bernard E. Leduc, Jehan H. Dagher, Pierre Mayer, François Bellemare and Yves Lepage

15. Low-Frequency Rectangular Pulse Is Superior to Middle Frequency Alternating Current Stimulation in Cycling of People With Spinal Cord Injury • ARTICLE
    Pages 338-345
    Johann Szecsi, Ché Fornusek, Phillip Krause and Andreas Straube
16. Effects of Osteoarthritis and Fatigue on Proprioception of the Knee Joint • ARTICLE
   Pages 346-350
   Meral Bayramoglu, Reyhan Toprak and Seyhan Sozay

17. Symmetry of Timing of Hip and Lumbopelvic Rotation Motion in 2 Different Subgroups of People With Low Back Pain • ARTICLE
   Pages 351-360
   Linda R. Van Dillen, Sara P. Gombatto, Dave R. Collins, Jack R. Engsberg and Shirley A. Sahrmann

18. Recovery of Gait After Short-Stay Total Hip Arthroplasty • ARTICLE
   Pages 361-367
   Inge van den Akker-Scheek, Martin Stevens, Sjoerd K. Bulstra, Johan W. Groothoff, Jim R. van Horn and Wiebren Zijlstra

19. Muscle Tone in Diabetic Polyneuropathy Evaluated by the Quantitative Pendulum Test • ARTICLE
   Pages 368-373
   Chou-Ching K. Lin, Ming-Shaung Ju and Han-Wei Huang

20. A Comparison of Psychometric Properties of the Smart Balance Master System and the Postural Assessment Scale for Stroke in People Who Have Had Mild Stroke • ARTICLE
   Pages 374-380
   Chi-Wen Chien, Ming-Hsia Hu, Pei-Fang Tang, Ching-Fan Sheu and Ching-Lin Hsieh

21. Ultrasound Imaging of Acute Biceps Tendon Changes After Wheelchair Sports • ARTICLE
   Pages 381-385
   Stefan van Drongelen, Michael L. Boninger, Bradley G. Impink and Tagreed Khalaf

Clinical Notes

22. Forced Use as a Potential Cause of Gastrocnemius Tears During Neurologic Rehabilitation: A Report of 2 Cases • ARTICLE
   Pages 386-388
   Steve R. Fisher, Laura L. Wiggs and Cindy B. Ivanhoe
23. The Complications of Scar Formation Associated With Intrathecal Pump Placement • SHORT COMMUNICATION
   Pages 389-390
   Marina G. Protopapas, Elizabeth Bundock, Susan Westmoreland, Christopher Nero, W. Andrew Graham and Shanker Nesathurai

   Pages 391-393
   Kenneth Mautner and John C. Keel

25. Ultrasonographic Findings of the Normal Ulnar Nerve in Adults • SHORT COMMUNICATION
   Pages 394-396
   Michael S. Cartwright, Hae W. Shin, Leah V. Passmore and Francis O. Walker

26. Exercise Interventions for Diabetes Control: Do We Really Know That Strength Training Is Better Than Endurance Training? • CORRESPONDENCE
   Page 397
   Douglas L. Weeks

27. The author responds • CORRESPONDENCE
   Pages 397-398
   Edmund Cauza

28. New Books • LITERATURE
   Page 399

29. Corrections • MISCELLANEOUS
   Page 400

30. Poster 11: The Acute Impact of Stroke on Functioning • ABSTRACT
   Page 400
   Lois E. Finch, J.M. Higgins and N.E. Mayo

31. Poster 15: Safety and Efficacy of Repeated Botulinum Toxin Type A Treatments for Focal Upper-Limb Poststroke Spasticity: Results of a 12-Month Multicenter, Open-Label Trial • ABSTRACT
   Page 400
   E. Elovic, A. Brashear, D. Kaelin, R. McIntosh, J. Liu and C. Turkel
Poster 78: Objective Assessment of Changes in Gait Using Pedobarography, While Walking and Running After Botulinum Toxin Injections in the Management of Equinovarus Deformity: A Case Study • ABSTRACT
Page 400
Gary N. Galang, Karen J. Nolan and Elie P. Elovic

Corrections • ERRATUM
Page 400

Corrections • ERRATUM
Page 400

Journal Based CME Evaluation and Application • MISCELLANEOUS
Page 401

CME Processing Fees and Application • MISCELLANEOUS
Page 402

ACRM News • NEWS
Pages 403-404
MISSION STATEMENT: The mission of the Archives of Physical Medicine and Rehabilitation is to disseminate information, with the ultimate goal of furthering the art and science of the practice of physical medicine and rehabilitation and interdisciplinary rehabilitation, and improving the health and welfare of persons with disabilities.
### ORIGINAL ARTICLES

<table>
<thead>
<tr>
<th>Page</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>273</td>
<td>A Randomized Controlled Trial of Modified Constraint-Induced Movement Therapy for Elderly Stroke Survivors: Changes in Motor Impairment, Daily Functioning, and Quality of Life</td>
<td>Ching-yi Wu, ScD, OTR, Chia-ling Chen, MD, PhD, Wen-chung Tsai, MD, PhD, Keh-chung Lin, ScD, OTR, Shih-han Chou, BS</td>
</tr>
<tr>
<td>279</td>
<td>Effect of Motorized Scooters on Physical Performance and Mobility: A Randomized Clinical Trial</td>
<td>Helen Hoenig, MD, MPH, Carl Pieper, DrPH, Laurence G. Branch, PhD, Harvey Jay Cohen, MD</td>
</tr>
<tr>
<td>287</td>
<td>Spasticity Experience Domains in Persons With Spinal Cord Injury</td>
<td>Jane S. Mahoney, DSN, RN, Joan C. Engebretson, DrPH, RN, Karen F. Cook, PhD, Karen A. Hart, PhD, Susan Robinson-Whelen, PhD, Arthur M. Sherwood, PhD</td>
</tr>
<tr>
<td>295</td>
<td>Motor Points for the Neuromuscular Blockade of the Subscapularis Muscle</td>
<td>Tim P. Harrison, MBBS, BSc, Anna Sadnicka, MBChB, BSc, Deborah M. Eastwood, MB, FRCS</td>
</tr>
<tr>
<td>298</td>
<td>Poststroke Shoulder Pain: Its Relationship to Motor Impairment, Activity Limitation, and Quality of Life</td>
<td>John Chae, MD, ME, Don Mascarenhas, MD, David T. Yu, MD, Andrew Kirsteins, MD, Elie P. Elovic, MD, Steven R. Flanagan, MD, Richard L. Harvey, MD, Richard D. Zorowitz, MD, Zi-Ping Fang, PhD</td>
</tr>
<tr>
<td>302</td>
<td>Measurement Properties of the National Institutes of Health Stroke Scale for People With Right- and Left-Hemisphere Lesions: Further Analysis of the Clomethiazole for Acute Stroke Study–Ischemic (Class-I) Trial</td>
<td>Scott R. Millis, PhD, Don Straube, MS, PT, Cherdsak Iramaneerat, MD, Everett V. Smith Jr, PhD, Patrick Lyden, MD</td>
</tr>
<tr>
<td>309</td>
<td>Clinimetric Properties of the Duruoz Hand Index in Patients With Stroke</td>
<td>Nebahat Sezer, MD, Gunes Yavuzer, MD, Koncuy Sivriglu, MD, Pinar Basaran, MD, B. Fusun Koseoglu, MD</td>
</tr>
<tr>
<td>315</td>
<td>Aerobic Capacity After Traumatic Brain Injury: Comparison With a Nondisabled Cohort</td>
<td>Kurt A. Mossberg, PhD, PT, Danielle Ayala, MPT, Tracey Baker, MPT, Justin Heard, MPT, Brent Masel, MD</td>
</tr>
<tr>
<td>321</td>
<td>The Development and Validity of the Salford Gait Tool: An Observation-Based Clinical Gait Assessment Tool</td>
<td>Brigite Toro, PhD, Christopher J. Nester, PhD, Pauline C. Farren, PhD</td>
</tr>
<tr>
<td>328</td>
<td>Inter- and Intraobserver Repeatability of the Salford Gait Tool: An Observation-Based Clinical Gait Assessment Tool</td>
<td>Brigite Toro, PhD, Christopher J. Nester, PhD, Pauline C. Farren, PhD</td>
</tr>
<tr>
<td>333</td>
<td>Estimated Prevalence of Obstructive Sleep Apnea–Hypopnea Syndrome After Cervical Cord Injury</td>
<td>Bernard E. Leduc, MD, FRCPC, Jehan H. Dagheer, MD, FRCPC, Pierre Mayer, MD, FRCPC, François Bellemare, PhD, Yves Lepage, PhD</td>
</tr>
<tr>
<td>338</td>
<td>Low-Frequency Rectangular Pulse Is Superior to Middle Frequency Alternating Current Stimulation in Cycling of People With Spinal Cord Injury</td>
<td>Johann Szecsi, MD, MSc (Eng), Che Formaseh, PhD, Phillip Krause, MD, Andreas Straube, MD</td>
</tr>
<tr>
<td>346</td>
<td>Effects of Osteoarthritis and Fatigue on Proprioception of the Knee Joint</td>
<td>Meral Bayramoglu, MD, Reyhan Toprak, MD, Seyhan Sozay, MD</td>
</tr>
</tbody>
</table>
### Table of Contents (continued)

<table>
<thead>
<tr>
<th>Page</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>351</td>
<td>Symmetry of Timing of Hip and Lumbopelvic Rotation Motion in 2 Different Subgroups of People With Low Back Pain</td>
<td>Linda R. Van Dillen, PhD, PT, Sara P. Gombatto, MSPT, Dave R. Collins, PhD, Jack R. Engsberg, PhD, Shirley A. Sahrmann, PhD, PT</td>
</tr>
<tr>
<td>361</td>
<td>Recovery of Gait After Short-Stay Total Hip Arthroplasty</td>
<td>Inge van den Akker-Scheek, MSc, Martin Stevens, PhD, Sjoerd K. Bulstra, PhD, MD, Johan W. Grootenhoff, PhD, Jim R. van Horn, PhD, MD, Wiebren Zijlstra, PhD</td>
</tr>
<tr>
<td>368</td>
<td>Muscle Tone in Diabetic Polyneuropathy Evaluated by the Quantitative Pendulum Test</td>
<td>Chou-Ching K. Lin, MD, PhD, Ming-Shaung Ju, PhD, Han-Wei Huang, MD</td>
</tr>
<tr>
<td>374</td>
<td>A Comparison of Psychometric Properties of the Smart Balance Master System and the Postural Assessment Scale for Stroke in People Who Have Had Mild Stroke</td>
<td>Chi-Wen Chien, BS, Ming-Hsia Hu, PhD, Pei-Fang Tang, PhD, Ching-Fan Sheu, PhD, Ching-Lin Hsieh, PhD</td>
</tr>
<tr>
<td>381</td>
<td>Ultrasound Imaging of Acute Biceps Tendon Changes After Wheelchair Sports</td>
<td>Stefan van Drongelen, PhD, Michael L. Boninger, MD, Bradley G. Impink, Tagreed Khalaf, MD</td>
</tr>
</tbody>
</table>

#### CLINICAL NOTES

<table>
<thead>
<tr>
<th>Page</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>386</td>
<td>Forced Use as a Potential Cause of Gastrocnemius Tears During Neurologic Rehabilitation: A Report of 2 Cases</td>
<td>Steve R. Fisher, MPT, Laura L. Wiggs, PT, Cindy B. Ivanhoe, MD</td>
</tr>
<tr>
<td>389</td>
<td>The Complications of Scar Formation Associated With Intrathecal Pump Placement</td>
<td>Marina G. Protopapas, DO, Elizabeth Bundock, MD, Susan Westmoreland, VMD, Christopher Nero, MD, W. Andrew Graham, PhD, Shanker Nesathurai, MD</td>
</tr>
<tr>
<td>391</td>
<td>Musculocutaneous Nerve Injury After Simulated Freefall in a Vertical Wind-Tunnel: A Case Report</td>
<td>Kenneth Mautner, MD, John C. Keel, MD</td>
</tr>
</tbody>
</table>

#### BRIEF REPORT

<table>
<thead>
<tr>
<th>Page</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>394</td>
<td>Ultrasonographic Findings of the Normal Ulnar Nerve in Adults</td>
<td>Michael S. Cartwright, MD, Hae W. Shin, MD, Leah V. Passmore, MS, Francis O. Walker, MD</td>
</tr>
</tbody>
</table>

#### DEPARTMENTS

<table>
<thead>
<tr>
<th>Page</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>397</td>
<td>Letter to the Editor</td>
<td>Exercise Interventions for Diabetes Control: Do We Really Know That Strength Training Is Better Than Endurance Training?</td>
</tr>
<tr>
<td>399</td>
<td>New Books</td>
<td></td>
</tr>
<tr>
<td>400</td>
<td>Corrections</td>
<td></td>
</tr>
</tbody>
</table>

#### ORGANIZATION NEWS—AAPM&R

<table>
<thead>
<tr>
<th>Page</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>401</td>
<td>Journal-Based CME Evaluation and Application</td>
<td></td>
</tr>
<tr>
<td>403</td>
<td>AAPM&amp;R 2007 Study Guide Included With This Issue of Archives</td>
<td></td>
</tr>
<tr>
<td>403</td>
<td>2007 AAPM&amp;R Officers Elected</td>
<td></td>
</tr>
<tr>
<td>403</td>
<td>Musculoskeletal Education and Research Grant Available From FPMR</td>
<td></td>
</tr>
<tr>
<td>403</td>
<td>AMA and AAPM&amp;R to Conduct Physician Practice Information Survey</td>
<td></td>
</tr>
<tr>
<td>404</td>
<td>Twenty-Five Year Members Honored by Academy</td>
<td></td>
</tr>
</tbody>
</table>
A Randomized Controlled Trial of Modified Constraint-Induced Movement Therapy for Elderly Stroke Survivors: Changes in Motor Impairment, Daily Functioning, and Quality of Life

Ching-yi Wu, ScD, OTR, Chia-ling Chen, MD, PhD, Wen-chung Tsai, MD, PhD, Keh-chung Lin, ScD, OTR, Shih-han Chou, BS

ABSTRACT. Wu C-Y, Chen C-L, Tsai W-C, Lin K-C, Chou S-H. A randomized controlled trial of modified constraint-induced movement therapy (mCIMT) on motor function, daily function, and health-related quality of life (HRQOL) in elderly stroke survivors.

Objective: To examine the benefits of modified constraint-induced movement therapy (mCIMT) on motor function, daily function, and health-related quality of life (HRQOL) in elderly stroke survivors.

Design: Two-group randomized controlled trial, with pre- and posttreatment measures.

Setting: Rehabilitation clinics.

Participants: Twenty-six elderly stroke patients (mean age, 72y) with 0.5 to 31 months postonset of a first-ever cerebrovascular accident.

Interventions: Twenty-six patients received either mCIMT (restraint of the unaffected limb combined with intensive training of the affected limb) or traditional rehabilitation for a period of 3 weeks.

Main Outcome Measures: Outcome measures included the Fugl-Meyer Assessment (FMA), FIM instrument, Motor Activity Log (MAL), and Stroke Impact Scale (SIS). The FMA evaluated the severity of motor impairment; the FIM instrument and MAL reported daily function; and the SIS detected HRQOL.

Results: The mCIMT group exhibited significantly greater improvements in motor function, daily function, and the physical domain of HRQOL than the traditional rehabilitation group. Patients in the mCIMT group perceived significantly greater percent of recovery after treatment than patients in the traditional rehabilitation group.

Conclusions: These findings suggest mCIMT is a promising intervention for improving motor function, daily function, and physical aspects of HRQOL in elderly patients with stroke.

The mCIMT was well tolerated by the elderly patients even though it is a rigorous training program.

Key Words: Controlled clinical trials; Occupational therapy; Quality of life; Rehabilitation; Stroke.

© 2007 by the American Congress of Rehabilitation Medicine and the American Academy of Physical Medicine and Rehabilitation

I T IS ESTIMATED THAT 75% of strokes occur in elderly patients. More than 50% of those 65 years and older who survive a stroke report persistent impairment of upper-extremity (UE) movement. They have been encouraged to use their unaffected UE to perform tasks and progressively avoid use of the affected UE during task performance. This behavior may result in learned nonuse phenomenon hindering a person’s recovery of movement and function in the affected limb. One approach that has shown great promise for enhancing UE motor performance and functional use of the affected UE among patients with stroke is constraint-induced movement therapy (CIMT). The specific techniques of CIMT involve restraining the use of the unaffected UE (6–20h/d for 2–3wk) and intense motor training (eg, 6h/d for 10–15 consecutive weekdays) through the use of shaping movements of the affected limb. The shaping procedure involves individualized task selection, graduated task difficulty, verbal feedback, prompting, and physically assisting with movements and modeling.

Although CIMT shows promise for improving motor deficits after stroke, converging data suggested that it may not be plausible in many environments. One possible reason is that intense and prolonged practice during CIMT may be less safe and more tiring particularly for elderly or deconditioned patients. To address the problems, Page et al devised a modified CIMT (mCIMT) with shorter training (eg, 2h/d on 10–15 consecutive weekdays) and restraint (eg, 6h/d for 2–3wk) time. The mCIMT program was shown to be applicable in chronic or subacute patients with a wide variety of motor disability and may be especially relevant for the elderly patients. However, no study has specifically examined the efficacy of mCIMT in elderly stroke survivors aged over 65.

A further gap in knowledge about the therapeutic benefits of mCIMT lies in the limited scope of outcome measures in prior research. Based on the International Classification of Functioning, Disability and Health framework and measurements at the impairment, activity, and participation levels may reflect the full range of domains affected by stroke. Nevertheless, mCIMT studies in general tend to capture effects on impairment level measures and activity level measures without evaluating outcomes of activity participation (eg, health-related quality of life [HRQOL]). The impairment level measure involved measures of synergy patterns, muscle strength, or motor

From the Graduate Institute of Clinical Behavioral Science and Department of Occupational Therapy, Chang Gung University, Tao-yuan, Taiwan (Wu, Chou); Department of Physical Medicine and Rehabilitation, Chang Gung Memorial Hospital, Tao-yuan, Taiwan (Chen, Tsai); School of Occupational Therapy, College of Medicine, National Taiwan University, and Department of Physical Medicine and Rehabilitation, National Taiwan University Hospital, Taipei, Taiwan (Lin). Supported by the National Health Research Institutes (grant no. NHRI-EX95-9103EC) and the National Science Council (grant no. NSC 93-2314-B-002-116). No commercial party having a direct financial interest in the results of the research supporting this article has or will confer a benefit upon the author(s) or upon any organization with which the author(s) is/are associated.
Correspondence to Keh-chung Lin, ScD, OTR, Sch of Occupational Therapy, College of Medicine, National Taiwan University and Department of Physical Medicine and Rehabilitation, National Taiwan University Hospital, 17, F4, Xu Zhou Rd, Taipei, Taiwan 100, e-mail: keclin@ha.mc.ntu.edu.tw; 0003-9993/07/8803-11134$32.00/0 doi:10.1016/j.apmr.2006.11.021
efficiency such as Fugl-Meyer Assessment (FMA), maximum grip strength,9,16 Wolf Motor Function Test (WMFT), and Nine Hole Peg Test.9,11,12,14 The activity level (or the functional level) measure involved assessment of performance of daily activities including objective measures such as FIM instrument and patient-oriented measures (eg, self-reported Motor Activity Log [MAL]).9,12,14 Functional performance after stroke may not correlate with level of HRQOL.17

One recent study18 investigated the effects of the CIMT on motor function (grip force for strength, Modified Ashworth Scale [MAS] for spasticity), daily function (WMFT and MAL for function in daily living), and HRQOL (Stroke Impact Scale [SIS]). As stated by these researchers, this study lacked a control therapy and the operation of other nonspecific effects cannot be ruled out. Cumulative data on comprehensive outcome measures for intervention effects are essential for evidence-based clinical decision making, research, and appropriate clinical management of stroke survivors.19 To examine the benefits of mCIMT in different aspects of health in elderly stroke survivors, we used FMA to reflect the improvement of motor function, FIM and MAL to objectively and subjectively represent daily function, and SIS to reflect HRQOL. This research used a randomized controlled trial to overcome the previous concern regarding research methodology in the study of Dettmers et al.18 The hypothesis was that patients receiving 3 weeks of mCIMT would exhibit substantially better performance in their affected UEs reflected by these 4 measures than patients receiving traditional rehabilitation.

METHODS

Participants

We recruited 26 elderly stroke patients (15 men, 11 women; mean age, 71.69y; range, 65–87y) from the rehabilitation departments of 3 medical centers and obtained informed consent. The subjects were right-handed before stroke by self-report, and were 0.5 to 31 months post onset of a first-ever cerebrovascular accident (mean, 7.5mo). To be included, the subject had to reach Brunnstrom stage III20 for the proximal part of UE or above. The other inclusion criteria were as follows: (1) no serious cognitive deficits (a score ≥63 on the modified Mini-Mental State Examination [MMSE]21); (2) considerable nonuse of the affected limb (an amount of use [AOU] score <2.5 on the MAL22); (3) no balance problems sufficient to compromise safety when wearing the project’s constraint device; and (4) no excessive spasticity in any of the joints of the affected UE (shoulder, elbow, wrist, fingers) (MAS score ≤2 in any joint).23 The institutional review board for human studies approved this protocol. All potential subjects received independent examinations by a physiatrist and an occupational therapist to determine their eligibility for inclusion.

Outcome Measures

We used FMA (maximum score, 66) to assess several dimensions of motor impairments by using a 3-point ordinal scale (0, cannot perform; 1, can perform partially; 2, can perform fully).24 Test-retest reliability, interrater reliability, and construct validity have been well established.25,26 The FIM instrument was used to objectively measure changes in activity performance through performance observation, and the MAL was used to subjectively measure changes in activity performance through self-report. The FIM instrument (maximum score, 126) consists of 18 items grouped into 6 subscales measuring self-care, sphincter control, transfers, locomotion, communication, and social cognition ability.27 Each item is rated with a score from 1 to 7 (1, complete assistance to perform basic activities of daily living [ADLs]; 2, maximal assistance; 3, moderate assistance; 4, minimal assistance; 5, supervision; 6, modified independence; 7, complete independence in performing basic ADLs). The FIM has established good interrater reliability.28–32

The MAL is a semistructured interview that obtained information about how patients use their affected limbs during 30 important ADLs. Patients used a 6-point AOU scale (score range, 0–5) to rate how much the arm is being used and a 6-point quality of movement (QOM) scale (score range, 0–5) to rate how well they are using their affected UEs.22

The SIS is a comprehensive measure of health outcomes in stroke populations. The evaluation of SIS involves a person’s participation in the activities that the person usually does in his/her life situation and relevant skills such as communication, memory, and mobility for participation in personal meaningful activities.33 The SIS was, thus, appropriate to be used to measure changes in HRQOL and participation performance through self-report.

The SIS, version 2, is a 64-item self-report scale designed to assess 8 functional domains including strength, memory, emotions, communication, ADLs and instrumental ADLs (IADLs), mobility, hand function, and participation, with established reliability and validity.33,34 Patients responded to items in each domain using a 5-point rating scale. Aggregated scores in each domain were generated and scores for each domain were computed using procedures published previously.33 A higher score means better performance.33

One question for assessing the patient’s global perception of percentage of recovery was included in the SIS. After patients finished the questions of the 8 domains, they were required to rate their percent recovery since their stroke on a visual analog scale of 0 to 100, with 0 indicating no recovery and 100 indicating full recovery.33

Design and Intervention

We applied a randomized pretest and posttest control group design. Subjects were individually randomized into the mCIMT or the traditional rehabilitation group by using a table of random numbers (fig 1).35 Before and after the 3-week intervention period, the tests were administered in random order by a blinded rater. Prior to administration of clinical measures (FMA, FIM), the blinded rater was trained to properly administer these 2 measures. This training included careful examination of written instructions and repeated practice. Rater competence was assessed by a senior certified occupational therapist.

For both groups, the study treatment occurred during the regularly scheduled occupational therapy (OT) session and all other routine interdisciplinary stroke rehabilitation proceeded as usual. When 2 or more study subjects were in the OT clinic at the same time, they were assigned to different treatment areas without opportunities to observe each other or rearrange to receive therapy at different times to prevent unintended crossover.

Each subject assigned to mCIMT participated in individualized, 2-hour therapy sessions, 5 times a week for 3 weeks. Shaping and adaptive and repetitive task practice techniques were used during the training sessions. Therapy concentrated on the affected limb use in functional tasks chosen by patients and the treating therapist, including turning on and off a light switch, reaching forward to move a jar from one place to another, picking up a cup and drinking from it, picking up a hairbrush and combing hair, and other activities similar to those performed on a daily basis. Approximately 15 minutes of
therapy was spent on normalization of muscle tone of the affected limb as needed. During the 3-week period, the patients’ unaffected hands and wrists were placed in mitts with self-adhesive (Velcro) straps every weekday for 6 hours identified as a time of frequent arm use.

With equivalent time and intensity of treatment, patients in the traditional rehabilitation group received standard therapy. During a 2-hour therapy session, approximately 75% of traditional rehabilitation focused on neurodevelopmental techniques emphasizing functional task practice when possible, as well as stretching of the affected limb, weight bearing with the affected limb, and fine motor dexterity activities. Approximately 25% of traditional rehabilitation focused on compensatory techniques using the unaffected limb to perform functional tasks and assist the affected limb during task performance.

Statistical Analysis

For all variables, we used analyses of covariance (ANCOVAs)\(^3\) to test whether, when controlling for pretreatment differences, the intervention improvement in the mCIMT group was greater than that in the traditional rehabilitation group. For each analysis, the pretest performance was the covariate, group was the independent variable, and posttest performance was the dependent variable. Effects sizes were calculated for each individual variable and indexed by using the effect size \(r\).\(^3\) According to Cohen, a large effect is represented by an \(r\) of at least .50, a moderate effect by .30, and a small effect by .10.\(^3\)

**RESULTS**

After being randomly assigned to 1 of the 2 groups, 13 subjects were included in the mCIMT group and 13 in the traditional rehabilitation group. The demographic and clinical characteristics of subjects in the 2 groups were comparable (table 1). Because the natural recovery of stroke patients with onset less than 6 months might be a confounder for study effects,\(^9\) we compared the onset time (mean onset: for mCIMT, 1.76mo; for traditional rehabilitation, 2.44mo) of patients whose onset was less than 6 months (mCIMT group, \(n=9\); traditional rehabilitation group, \(n=8\)) between the 2 groups and found no significant differences (\(P=.385\)) between the groups.

Table 2 shows the descriptive statistics for each outcome measure. The SIS score (\(P=.039\)) and the scores of participation (\(P=.004\)) and perceived recovery (\(P=.012\)) domains of the SIS showed significant differences between the groups at pretreatment time point. No other outcome measures showed significant differences between the groups before treatment.

---

**Table 1: Characteristics of Study Participants**

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>mCIMT ((n=13))</th>
<th>TR ((n=13))</th>
<th>(P^*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex (male/female)</td>
<td>8/5</td>
<td>7/6</td>
<td>.691</td>
</tr>
<tr>
<td>Age (y)</td>
<td>71.44±6.42</td>
<td>71.94±6.79</td>
<td>.849</td>
</tr>
<tr>
<td>Side of lesion (right/left)</td>
<td>6/7</td>
<td>7/6</td>
<td>.695</td>
</tr>
<tr>
<td>Months after stroke</td>
<td>6.70±8.99</td>
<td>8.32±7.97</td>
<td>.616</td>
</tr>
<tr>
<td>Brunnstrom stage of proximal part of UE (median)</td>
<td>4.5</td>
<td>4.5</td>
<td>.312</td>
</tr>
<tr>
<td>Modified MMSE</td>
<td>82.92±11.21</td>
<td>83.08±11.41</td>
<td>.973</td>
</tr>
<tr>
<td>AOU MAL</td>
<td>0.80±1.38</td>
<td>1.37±1.71</td>
<td>.356</td>
</tr>
<tr>
<td>MAS</td>
<td>0.26±0.32</td>
<td>0.32±0.34</td>
<td>.588</td>
</tr>
</tbody>
</table>

NOTE: Values are mean ± standard deviation (SD) or as otherwise indicated.
Abbreviation: TR, traditional rehabilitation.
\(^*P\) associated with the chi-square test for categorical variables, with the independent \(t\) test for continuous variables, and with the Mann-Whitney \(U\) test for ordinal variables.
NOTE. Values are mean ± SD.

Table 3 shows the results of the ANCOVAs that tested the effects of mCIMT relative to traditional rehabilitation. The results showed significant and moderate-to-large effects in favor of the mCIMT group on FMA, FIM, and MAL. Patients in the mCIMT group reported greater improvements in AOU and in QOM of their affected limbs during daily activities. The mCIMT group reported using the affected UE for an average of 14 activities before treatment, and the mCIMT for 24 and the traditional rehabilitation for 22 after treatment.

The results also showed significant and moderate effects of mCIMT on the overall SIS and some aspects of the SIS. These greater improvements were shown in a few QOL domains on the SIS including strength and ADLs and IADLs. There were greater improvements were shown in a few QOL domains on the SIS including strength and ADLs and IADLs. There were substantially lower scores of FMA, FIM, and MAL before treatment were slightly lower in the mCIMT group than in the traditional rehabilitation group.

The substantial improvement in the abnormal movement patterns, reflected by FMA, in the mCIMT group suggested that mCIMT reversed impairments rather than simply helped patients to adapt to residual impairments. Accordingly, largely improved daily function, reflected by FIM, in the mCIMT group may result from the reduced motor impairments rather than developing new compensatory strategies. The score changes in FIM were supported by the improvement reported for the ADL and IADL domain of the SIS.

The greater improvement in the scores of the FMA and the FIM seen in the mCIMT group than in the traditional rehabilitation group corresponded with those of previous studies.9,11,12,14 The substantial improvement in the abnormal movement patterns, reflected by FMA, in the mCIMT group suggested that mCIMT reversed impairments rather than simply helped patients to adapt to residual impairments. Accordingly, largely improved daily function, reflected by FIM, in the mCIMT group may result from the reduced motor impairments rather than developing new compensatory strategies. The score changes in FIM were supported by the improvement reported for the ADL and IADL domain of the SIS.

Patients in the mCIMT group subjectively reported considerably larger improvements in the use and function of their affected UEs, as measured by the MAL, than those in the traditional rehabilitation group. These findings on MAL are consistent with previous findings.9,11,12,14 These MAL scores in the mCIMT group suggested that the learned nonuse phenomenon observed in the patients can be overcome through a modified intensive training and mitral wear schedule emphasizing repeated functional use.

As shown in the descriptive statistics (table 2), the mean scores of FMA, FIM, and MAL before treatment were slightly lower in the mCIMT group than in the traditional rehabilitation group, though nonsignificant differences were found. However, the mean scores of these 3 measures after treatment were higher in the mCIMT group than in the traditional rehabilitation group, demonstrating that patients receiving mCIMT exhibited improvements in reduced motor impairment and enhanced functional use of the affected UE in daily activities.

DISCUSSION

This randomized controlled study supported in an elderly sample the effectiveness of mCIMT in stroke patients. Patients improved in different aspects of motor function, daily function, and participation as reflected by the UE movement patterns, independence in ADLs, and some aspects of QOL. There was no attrition and full protocol adherence, indicating that the mCIMT is well tolerated for the elderly stroke patients.

The greater improvement in the scores of the FMA and the FIM seen in the mCIMT group than in the traditional rehabilitation group corresponded with those of previous studies.9,11,12,14 The substantial improvement in the abnormal movement patterns, reflected by FMA, in the mCIMT group suggested that mCIMT reversed impairments rather than simply helped patients to adapt to residual impairments. Accordingly, largely improved daily function, reflected by FIM, in the mCIMT group may result from the reduced motor impairments rather than developing new compensatory strategies. The score changes in FIM were supported by the improvement reported for the ADL and IADL domain of the SIS.

Patients in the mCIMT group subjectively reported considerably larger improvements in the use and function of their affected UEs, as measured by the MAL, than those in the traditional rehabilitation group. These findings on MAL are consistent with previous findings.9,11,12,14 These MAL scores in the mCIMT group suggested that the learned nonuse phenomenon observed in the patients can be overcome through a modified intensive training and mitral wear schedule emphasizing repeated functional use.

As shown in the descriptive statistics (table 2), the mean scores of FMA, FIM, and MAL before treatment were slightly lower in the mCIMT group than in the traditional rehabilitation group, though nonsignificant differences were found. However, the mean scores of these 3 measures after treatment were higher in the mCIMT group than in the traditional rehabilitation group, demonstrating that patients receiving mCIMT exhibited improvements in reduced motor impairment and enhanced functional use of the affected UE in daily activities.

Table 2: Performance on the Outcome Measures From Pre- to Post-Treatment

<table>
<thead>
<tr>
<th>Measures</th>
<th>Pretreatment (Mean ± SD)</th>
<th>Post-Treatment (Mean ± SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FMA (UE)</td>
<td>41.85 ± 13.33</td>
<td>49.54 ± 12.84</td>
</tr>
<tr>
<td>FIM</td>
<td>95.08 ± 15.24</td>
<td>104.85 ± 12.13</td>
</tr>
<tr>
<td>MAL</td>
<td>53.13 ± 8.95</td>
<td>62.22 ± 8.71</td>
</tr>
<tr>
<td>AOU</td>
<td>0.80 ± 1.38</td>
<td>1.78 ± 1.28</td>
</tr>
<tr>
<td>QOM</td>
<td>0.79 ± 1.29</td>
<td>1.99 ± 1.31</td>
</tr>
<tr>
<td>SIS</td>
<td>53.13 ± 8.95</td>
<td>62.22 ± 8.71</td>
</tr>
</tbody>
</table>

Table 3: Results of Inferential Statistics on the Outcome Measures

<table>
<thead>
<tr>
<th>Measures</th>
<th>F_{1,22}</th>
<th>P</th>
<th>Effect Size r</th>
</tr>
</thead>
<tbody>
<tr>
<td>FMA (UE)</td>
<td>6.87</td>
<td>.008*</td>
<td>.48</td>
</tr>
<tr>
<td>FIM</td>
<td>4.94</td>
<td>.018*</td>
<td>.42</td>
</tr>
<tr>
<td>MAL</td>
<td>9.75</td>
<td>.003*</td>
<td>.55</td>
</tr>
<tr>
<td>AOU</td>
<td>14.76</td>
<td>&lt;.001*</td>
<td>.63</td>
</tr>
<tr>
<td>QOM</td>
<td>3.29</td>
<td>.042*</td>
<td>.35</td>
</tr>
<tr>
<td>SIS</td>
<td>5.33</td>
<td>.015*</td>
<td>.43</td>
</tr>
<tr>
<td>Hand function</td>
<td>0.71</td>
<td>.204</td>
<td>.17</td>
</tr>
<tr>
<td>Mobility</td>
<td>1.63</td>
<td>.107</td>
<td>.26</td>
</tr>
<tr>
<td>ADLs/IADLs</td>
<td>3.12</td>
<td>.045*</td>
<td>.35</td>
</tr>
<tr>
<td>Memory</td>
<td>0.02</td>
<td>.444</td>
<td>.03</td>
</tr>
<tr>
<td>Communication</td>
<td>0.59</td>
<td>.226</td>
<td>.16</td>
</tr>
<tr>
<td>Emotion</td>
<td>0.01</td>
<td>.464</td>
<td>.00</td>
</tr>
<tr>
<td>Participation</td>
<td>0.98</td>
<td>.167</td>
<td>.20</td>
</tr>
<tr>
<td>Stroke recovery</td>
<td>13.36</td>
<td>.001*</td>
<td>.59</td>
</tr>
</tbody>
</table>

*P<.05.
The possible mechanisms responsible for improvement in motor function and daily function after such a short period of therapy can be speculated, based on the literature. First, reinforcement of using the affected limb and aversive consequences for its nonuse by constraining the unaffected hand may reduce the learned nonuse behavior. Second, repeated practice of functional tasks of ecologic significance to the patients may lead to increased reorganization of the brain after stroke. This use-dependent cortical reorganization may represent the neural basis of increased use of the affected UE. The natural recovery of stroke patients with onset less than 6 months might be a confounder for the explanation of the beneficial effects of mCIMT. However, the numbers of subjects with onset less than 6 months between the 2 groups were similar and the difference in mean onset time was nonsignificant. Thus the observed effects in favor of mCIMT cannot be attributed to the confounding effect of natural recovery after stroke.

The mCIMT group appeared to obtain greater gains in HRQOL than traditional rehabilitation. These gains were mostly shown in the physical domains (ie, strength and ADLs and IADLs), which is partially consistent with the previous study. Because the mCIMT program targeted functional training of movement, patients should directly obtain gains in the physical performance and daily function. It should also be noted that the traditional rehabilitation group showed lower scores in some domains of HRQOL after treatment. The posttreatment differences between the study groups are due, in part, to the posttreatment declines in the traditional rehabilitation group.

Patients receiving mCIMT did not subjectively exhibit significantly greater improvements in hand function than those receiving traditional rehabilitation, inconsistent with the findings of the previous study. The possible reason is that hand dexterity and perceptual-motor adaptability decreases with age, especially after the age of 65 years. The potential for relearning hand function through rehabilitation training is, thus, limited and no further improvement was found in mCIMT. Ranganathan et al suggested that skilled finger movement exercise such as holding 2 metal balls in the palm of the hand and rotating the balls smoothly clockwise or counterclockwise improves hand function in elderly people. Future research may investigate whether incorporating intensive exercise of skilled finger movements into mCIMT may improve subjective perceptions of hand function.

Similar to the findings of previous research, patients in the mCIMT group did not perceive further improvement in mobility than those in the traditional rehabilitation group possibly because the training program did not involve transfer or mobility tasks. No significant differences after treatment between 2 groups on the memory and thinking, emotion, communication, and participation domains suggested that the effects of intensive physical training such as mCIMT on physical performance may not generalize into the effects on cognitive and psychosocial domains. Because mCIMT is an intensive training during 2 to 3 weeks, the interaction between therapists and patients is an important component of this treatment. To enhance the psychosocial well-being of individual patients, the therapist should consider patients’ opinions regarding how the functioning gained from mCIMT applies to their social lives. mCIMT programs that link motor function to social activities may enhance social participation more effectively. Examples of such treatments are those that involve dyadic and group interaction on collaborative activities and practice on constituent tasks relevant for community function (eg, simulated activities of grocery shopping). Furthermore, home-based mCIMT might also be appropriate for use in the elderly to facilitate generalization of therapeutic gains in motor function and self-care skills to daily life and community function. Considerations of the practice context and activity parameters are especially relevant for mCIMT to be implemented with success in the aged populations. Future research may study whether mCIMT using client-valued activities for task-oriented practice in the home setting or domicile community would be more beneficial than hospital-based physical training for improving function and preventing disability in the elderly.

Study Limitations

A few limitations to this study warrant consideration. First, the treatment effects were measured immediately after treatment and the beneficial effects of mCIMT may enhance social participation more effectively. Examining during 2 to 3 weeks, the interaction between therapists and patients may have resulted in nonequivalency in some of the outcome measures such as the participation and perceived recovery aspects of the SIS. To correct for this problem, ANCOVAs were used to control for the pretreatment differences between groups. Future research using stratified random sampling (eg, matching groups on baseline characteristics) might serve to better control for the problem of pretreatment differences between groups. A final limitation pertains to a problem that is characteristic of all CIMT studies. Although all subjects received the same intensity and duration of treatment intervention, the CIMT group arguably received more “treatment” during restraint wear out of clinic. Future research may use a control group that receives traditional rehabilitation together with restraint wear out of clinic to address this potential bias.

Additional considerations for extended research include measurements of functional independence in various performance contexts (eg, hospital-based measures of self-care and mobility and evaluation for IADLs after hospital discharge). Future research may also study factors that may affect treatment outcomes (eg, stroke severity, side of hemiplegia, motivation for treatment participation). Such research may reveal prognostic factors relevant for outcome prediction and patient selection.

CONCLUSIONS

The unique contribution of this study is to investigate the feasibility and efficacy of mCIMT for improving affected limb use, daily function, and HRQOL in elderly stroke survivors. The findings suggest that mCIMT improves movement performance and ADL abilities as measured by clinical tests, whether subjective or objective, mCIMT improves physical aspects of HRQOL and was well tolerated by the elderly patients although it is a rigorous training program. Future clinical trials may enroll a larger sample for follow-up study to evaluate the long-term benefits of mCIMT in the elderly. More focused evaluation of rehabilitation practice that aims at improving aspects of HRQOL and functional domains at the activity level in the elderly is needed. This treatment evaluation will contribute to improved practice for elderly stroke survivors.

References

Effect of Motorized Scooters on Physical Performance and Mobility: A Randomized Clinical Trial

Helen Hoenig, MD, MPH, Carl Pieper, DrPH, Laurence G. Branch, PhD, Harvey Jay Cohen, MD


Objective: To investigate the effects of providing a motorized scooter on physical performance and mobility.

Design: Randomized clinical trial comparing scooter users with usual care.

Setting: One academic and 1 Veterans Affairs medical center.

Participants: Ambulatory, community-dwelling outpatients with rheumatoid arthritis or osteoarthritis of the knee.

Intervention: Provision of a motorized scooter for 3 months.

Main Outcome Measures: Six-minute walk distance (6MWD) and mobility methods in diverse locations at baseline, 1 month, and 3 months, and accidents while using the scooter.

Results: The majority of scooter subjects (n = 16/22 [72.7%]) used the scooter 4 or more days per week. The difference ± standard deviation between the 2 groups in change in 6MWD over the study period was not statistically significant (scooter users, 16.9 ± 7.3 m; usual care, 17.2 ± 7.5 m; P = .55). Four (18.1%) scooter users reported accidents. Over the study period, the proportion of persons reporting use of a scooter (provided by the study or otherwise available) increased in the scooter-users group (eg, food stores, 16.7% to 52.6%; doctor’s office, 0% to 9.1%).

Conclusion: Motorized scooters provided to ambulatory persons with arthritis were used intermittently. The greatest short-term risk from scooter usage appeared to be minor collisions.

Key Words: Activities of daily living; Assistive technology; Bedrest; Cardiovascular deconditioning; Duration of medical equipment; Exercise; Mobility limitation; Occupational therapy; Osteoarthritis; Outcome and process assessment; Physical therapy; Randomized controlled trials; Rehabilitation; Residential mobility; Rheumatoid arthritis; Walking; Wheelchairs.

Third-party payers require physician prescription of durable medical equipment such as motorized scooters, necessitating physician judgments about their medical appropriateness. Although the usage of tools to cope with mobility impairments dates back at least 3000 years, data are limited on the effectiveness of mobility equipment as it is applied to daily life.

Scooter prescription, especially for ambulatory patients, is particularly controversial because of the expense of the device, recent Medicare fraud, and uncertainty about the risk of deconditioning. Given the compelling data supporting the health benefits of physical activity, a scooter might cause harm if it were used instead of walking. Alternatively, some argue that a motorized scooter might increase activity levels by enabling people to participate in activities that would otherwise not be possible. The lack of empirical data leaves uncertainty about the clinical indications for prescription of a motorized scooter.

The overarching purpose of this study was to study objectively the effects of provision of a mobility aid on physical performance and day-to-day mobility. Motorized scooters were used as the exemplar device because, as discussed earlier, scooters are particularly controversial and their usage is more likely to be optional than other types of mobility aids. The primary research question was whether provision of a motorized scooter for 3 months would result in a clinically significant change in 6-minute walk distance (6MWD) compared with similar subjects who did not have a motorized scooter.

Secondary questions included the effect of provision of a motorized scooter on mobility in diverse community locations, accidents related to scooter use, and subjective perceptions of the helpfulness of a scooter in daily life.

METHODS

Approval was obtained for this study from the institutional review boards of Duke University and the Durham Veterans Affairs (VA) Medical Center (VAMC), and all subjects provided informed consent.

Study Design and Intervention

This was a randomized trial performed by using a computer-generated random numbers table and concealed allocation (sequential, sealed envelopes filled by a person unconnected with the study), administered by the project coordinator after subject enrollment. Neither the subjects nor the investigators were blinded to the intervention.

The goal of the study was to examine the effectiveness of a motorized scooter. Effectiveness studies examine the “usefulness of a particular treatment to the individuals receiving it under typical clinical conditions.” The intervention con-
sisted of provision of a motorized scooter (the Legend®) and a lift to transport the scooter (the Backsaver® or the Outrider®) for a 3-month period of time. Study personnel did not provide subjects with directions on when or where to use the scooter. Per standard clinical practices at the time of the study, subjects were instructed in the operation of the scooter and lift by the medical supplier at the time the lift was installed on the car. All subjects and their physicians were instructed to continue any regular exercise program and medications during the study period but to avoid starting or changing medications or exercises, including walking specifically for exercise.

Patient Sample and Setting

Several considerations determined the study population. More physically impaired, activity-restricted persons might use the scooter more, yet those same people might have less potential to experience deconditioning because of already being deconditioned. Validated measures of physical performance for physically impaired persons are limited in so far as they may require mobility skills out of the range of persons with disability and/or have inaccurate results because of altered physiologic responses and/or lack validation.10 Therefore, the study sample was limited to adults who were able to walk independently for at least 15m (50ft) based on direct observation over a 15-m course. To control for disease-specific effects, enrollment was limited to persons who met American Rheumatological Association (ARA) clinical criteria for osteoarthritis (OA) or the usual care.

Study sample was limited to adults who were able to walk independently for at least 15m (50ft) based on direct observation over a 15-m course. To control for disease-specific effects, enrollment was limited to persons who met American Rheumatological Association (ARA) clinical criteria for osteoarthritis (OA) or rheumatoid arthritis (RA)11 based on physical examination and review of the medical record. Other inclusion and exclusion criteria included approval by the patient’s primary care physician or rheumatologist, no cardiac disease that would preclude exercise testing (eg, recent acute myocardial infarction), stable rheumatologic disease (<10mg/d of prednisone, no change in medications in the preceding 3 months, no surgery planned for the next 3–6 months), concurrence to avoid changes in medications for the next 3 to 6 months, a valid North Carolina driver’s license, a Mini-Mental State Examination (MMSE)12 score of greater than 26/30, and ownership of an automobile that would accommodate a scooter lift and heavy-duty shock absorbers.

Study subjects were recruited from January 1999 through December 2002 by using methods that might be typical of commercial vendors (advertisements and directly mailing), including advertisements in the newspaper and in 2 medical centers (VA, 1 academic) and direct mailing (4586 letters) to patients from an academic rheumatologic practice, patients from the Rheumatology and Orthopedic Clinics at the VAMC, and patients at the VA Ambulatory Care Clinic or discharged from the VA hospital with a primary diagnosis of OA or RA (International Classification of Diseases, 9th Revision, Clinical Modification [ICD-9-CM], codes 714.X–715.X) or orthopedic knee surgery (ICD-9-CM procedure codes 81.22, 81.47, 80.26, 80.16, 81.54). Subjects called a central number and were screened over the telephone via self-report to affirm a diagnosis of RA or OA with knee pain, ability to walk, and type of automobile. Persons who screened positive telephonically were seen in person to obtain the MMSE and informed consent; then, a physician or physical therapist checked ARA criteria for OA or RA and the ability to walk; and, finally, the patient’s physician was contacted for approval. If the person met all eligibility criteria, he/she was randomized to receive a scooter or the usual care.

The study had a limited number of motorized scooters, constraining the number of active subjects at any given time, although subjects might be screened and randomized at any time. If the subject was in the scooter arm of the study, baseline data were collected within 1 week before modifying the car and training in use of the scooter. The median length of time from randomization to baseline data collection was 36 days (scooter, 41; usual care, 29), and the mean was 47 days (scooter, 43; usual care, 50.3).

Measures

All measures, except self-reported measures on scooter accidents and satisfaction, were collected in person, on site at the VAMC at baseline (1-wk before provision of the scooter), 1 month, and 3 months.

Dependent Variables

Six-minute walk distance. The 6MWD was selected as the primary outcome measure for several reasons. The 6MWD correlates highly with other measures of functional performance, test-retest reliability is excellent,14 it has been shown to be sensitive to the effects of walking exercise in subjects with knee OA,15–17 and it appears to be more sensitive to the effects of walking exercise than other physical performance measures.18

Measuring assistive technology outcomes can be challenging.19,20 Nonetheless, several additional outcomes were examined, including mobility, accidents, and satisfaction.

Mobility. Based on a previously developed questionnaire,21 we collected self-reported data at 1 and 3 months on the mobility method (walked without assistance, walked with assistance, used scooter, other) used during the preceding week in a variety of locations (work place, school, friend or relative’s house, food store, drug store, other store, restaurant or bar, religious building, park, movie, sporting event, library, other places of recreation, civic building, bank or ATM [asynchronous transfer mode] machines, doctor’s office, dentist or other medical office, other). Results are reported as the proportion of subjects going to a specific location, the proportion using a scooter among those who went to that location, and a count of the total number of different locations visited. Data were collected on number of steps per day for a 1-week period at baseline and at 1 and 3 months by using a pedometer.22

Scooter accidents. At 1 month and 3 months, subjects reported if they had experienced an “accident” while using the scooter, and, if so, how many accidents they had experienced and they described the accidents.

Satisfaction. At 3 months, scooter users reported how helpful they found the scooter (5-level Likert scale ranging from very helpful to very unhelpful), if they intended to obtain a scooter of their own (yes vs no), and they answered several open-ended questions including (1) “In what ways was the scooter helpful?” (2) “In what ways was the scooter unhelpful?” and (3) “In what ways did you use the scooter?”

Independent Variables

Because of the relatively small sample size and the randomized methodology, the only independent variable included in the primary analysis was study group assignment. Several additional variables were measured for purposes of secondary analyses and to describe the population.

Disease characteristics. Disease activity was measured by pain. Many measures of disease activity in RA are not applicable to OA. Pain, on the other hand, is responsive to therapeutic interventions in RA22,23 and to the effects of walking exercise in patients with OA.24 and it might directly affect the primary outcome measure, 6MWD. Pain was measured with a 10-cm, horizontal, double-anchored visual analog scale at baseline and at 3 months.25
Physical function. Physical function (difficulty with self-care tasks and use of assistive technology for mobility in particular) was measured with the Health Assessment Questionnaire and by measurement of the number of seconds to walk 15m at a comfortable walking pace.

Sociodemographic characteristics. Sociodemographic variables include self-reported data on date of birth, sex, race, and education.

Intervention characteristics. Frequency of scooter usage was measured by self-report at 3 months by using a Likert scale (more than once daily, once daily, 4–6 times a week, 2–3 times a week, about once a week, less than once a week). The question structure was derived from questionnaires of similar types of activities and the response options based on typical wheelchair use found in prior work. Studies have shown self-report and objective measures of physical activity to have moderate to strong correlations. Prior studies showed that self-reported wheelchair usage did not suffer from recall bias and it had good construct validity (eg, environmental barriers predicted lower self-reported wheelchair usage).

Statistical Analysis

The primary outcome was change in 6MWD over 3 months; thus, analyses were limited to subjects with data at baseline and at 3 months. Because subjects were randomized to a group, the primary analysis was a t test (and the nonparametric equivalent to the t test, the 2-sample Wilcoxon). Follow-up sensitivity analyses explored whether the observed effect differed according to baseline walking speed (gait speed above vs below the median) or pain at the end of the trial (above vs below the median) by using a general linear model with group, the predictor, and the interaction of these 2 variables.

Power. The study goal was to enroll 50 persons, with an estimated dropout rate of 20% for a final sample of 40 persons completing the study, which would provide the statistical power to detect a standardized difference of approximately .80 in 6MWD, with a 2-sided α of .05 and β of .20, providing sufficient power to detect the moderate to large effect size that might be expected with a substantive change in walking activity because of the scooter and distances equivalent to those required for community mobility. The final study sample of 43 subjects provided sufficient power to detect a standardized difference of .86 (ie, a change of ≥59.7m [≥199ft] in 6MWD based on the standard deviation [SD] of this study sample) with a 2-sided α of .05 and β of .20.

RESULTS

Fifty-three persons met inclusion and exclusion criteria and were randomized (scooter, 26; usual care, 26). Forty-seven started the study (scooter, 24; usual care, 23), and 43 completed the study (scooter, 22; usual care, 21) (fig 1). Knee OA was the primary diagnosis in most subjects enrolled in the study (scooter, 19; usual care, 21), the balance having RA. The study population was reflective of the veteran population in general (table 1), with a mean age of 63 years, 60% white, 79% male, and 26% with less than a high school education. The majority of participants (67%) reported that they already were using a mobility aid, primarily a cane. Over 60% of participants reported difficulty walking on level ground, and 86% reported difficulty climbing stairs. During the week before starting the study, the mean pain level was 5.5/10. Statistically significant differences at baseline between the scooter and the usual-care groups were present for age and for wheelchair usage (although the 2 groups did not differ in use of other mobility aids). There were few changes in medication during the study (medications changes: scooter, 4; usual care, 5). On average, the scooters were used more than a high school education. The majority of participants (67%) reported that they already were using a mobility aid, primarily a cane. Over 60% of participants reported difficulty walking on level ground, and 86% reported difficulty climbing stairs. During the week before starting the study, the mean pain level was 5.5/10. Statistically significant differences at baseline between the scooter and the usual-care groups were present for age and for wheelchair usage (although the 2 groups did not differ in use of other mobility aids). There were few changes in medication during the study (medications changes: scooter, 4; usual care, 5). On average, the scooters were used several times weekly, with 40.9% of the sample reporting daily scooter usage.

Table 1: Sample Demographic and Clinical Characteristics for the 2 Study Groups

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Scooter (n=22)</th>
<th>Usual Care (n=21)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age ± SD (y)</td>
<td>67.2±9.4</td>
<td>58.2±11.7</td>
<td>.01</td>
</tr>
<tr>
<td>Sex, male (%)</td>
<td>86.4</td>
<td>71.4</td>
<td>.24</td>
</tr>
<tr>
<td>Race, white (%)</td>
<td>72.3</td>
<td>52.4</td>
<td>.18</td>
</tr>
<tr>
<td>Education, less than high school (%)</td>
<td>18.2</td>
<td>33.3</td>
<td>.26</td>
</tr>
<tr>
<td>Mean hours out of bed baseline ± SD</td>
<td>13.9±3.1</td>
<td>14.5±3.6</td>
<td>.51</td>
</tr>
<tr>
<td>Mobility aid use, any vs none (%)</td>
<td>59.1</td>
<td>76.2</td>
<td>.24</td>
</tr>
<tr>
<td>Specific mobility aids used (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cane</td>
<td>54.5</td>
<td>66.7</td>
<td>.42</td>
</tr>
<tr>
<td>Walker</td>
<td>9.1</td>
<td>9.5</td>
<td>.96</td>
</tr>
<tr>
<td>Crutches</td>
<td>13.6</td>
<td>14.3</td>
<td>.95</td>
</tr>
<tr>
<td>Wheelchair</td>
<td>18.2</td>
<td>0</td>
<td>.04</td>
</tr>
<tr>
<td>Mean pain, baseline ± SD</td>
<td>5.2±2.2</td>
<td>5.8±2.5</td>
<td>.53</td>
</tr>
<tr>
<td>Difficulty walking on level ground (%)</td>
<td>63.6</td>
<td>66.7</td>
<td>.85</td>
</tr>
<tr>
<td>Difficulty climbing stairs (%)</td>
<td>90.9</td>
<td>80.9</td>
<td>.11</td>
</tr>
</tbody>
</table>
There was no statistically significant difference between the 2 groups ($P=.55$) in change in 6MWD over time (Table 2). Supplemental exploratory analyses (Table 3) did not reveal a decline in 6MWD among particular subgroups of scooter users, nor were there statistically significant differences in 6MWD compared with usual care among the particular subgroups examined. Scooter subjects with worse gait at baseline had a strong trend ($P=.10$) for increased 6MWD over the study period compared with usual care; however, this was because the faster walkers in the usual-care group actually declined in 6MWD. Pain above versus below the median at 3 months did not relate to change in 6MWD ($P=.75$). 6MWD increased more in daily scooter users compared with those with less frequent scooter usage, but this did not reach statistical significance ($P=.39$).

Table 4 shows data on mobility for the 4 most commonly accessed locations: shopping, visiting friends, going to the doctor, and going to church. Although there was little change in the proportion of subjects going to specific locations, there was a clear increase in the scooter group in use of a scooter particularly in food stores and going to the doctor. In response to an open-ended question on the ways the scooter was used, there was little change in the proportion of subjects going to specific locations, there was a clear increase in the scooter group in use of a scooter particularly in food stores and going to the doctor.

Table 2: Change in 6MWD Over the Study Period

<table>
<thead>
<tr>
<th>Change in 6MWD</th>
<th>Scooter</th>
<th>Usual Care</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline to 1 month (m)</td>
<td>$16.6 \pm 71.1$ (54.4 $\pm$ 235.2ft)</td>
<td>$18.2 \pm 53.5$ (59.6 $\pm$ 175.5ft)</td>
<td>.66</td>
</tr>
<tr>
<td>Baseline to 3 months (m)</td>
<td>$16.9 \pm 73.0$ (55.5 $\pm$ 239.6ft)</td>
<td>$17.2 \pm 69.5$ (56.5 $\pm$ 228.0ft)</td>
<td>.55</td>
</tr>
</tbody>
</table>

NOTE. Values are mean $\pm$ SD.

Table 3: Exploratory Analyses of Change in 6MWD Over the 3-Month Study Period for Selected Subgroups

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>Scooter</th>
<th>Usual Care</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline gait speed</td>
<td></td>
<td></td>
<td>.10</td>
</tr>
<tr>
<td>Slow ($&lt;$median), m</td>
<td>$5.3 \pm 61.1$ (17.4 $\pm$ 200.4ft)</td>
<td>$42.6 \pm 60.3$ (139.8 $\pm$ 197.9ft)</td>
<td></td>
</tr>
<tr>
<td>Fast ($&gt;$median), m</td>
<td>$28.5 \pm 84.7$ (93.5 $\pm$ 277.8ft)</td>
<td>$8.1 \pm 71.6$ (26.7 $\pm$ 234.8ft)</td>
<td></td>
</tr>
<tr>
<td>Baseline pain</td>
<td></td>
<td></td>
<td>.75</td>
</tr>
<tr>
<td>Less pain ($&lt;$median), m</td>
<td>$20.1 \pm 88.0$ (65.9 $\pm$ 288.7ft)</td>
<td>$12.6 \pm 86.6$ (41.4 $\pm$ 284.1ft)</td>
<td></td>
</tr>
<tr>
<td>More pain ($&gt;$median), m</td>
<td>$21.0 \pm 63.1$ (68.8 $\pm$ 207.0ft)</td>
<td>$23.1 \pm 45.0$ (75.5 $\pm$ 147.6ft)</td>
<td></td>
</tr>
<tr>
<td>Frequency of scooter usage</td>
<td></td>
<td></td>
<td>.39</td>
</tr>
<tr>
<td>Daily usage, m</td>
<td>$33.4 \pm 64.4$ (109.4 $\pm$ 211.2ft)</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Less than daily usage, m</td>
<td>$5.52 \pm 85.2$ (18.1 $\pm$ 279.6ft)</td>
<td>NA</td>
<td></td>
</tr>
</tbody>
</table>

NOTE. Values are mean $\pm$ SD.

Abbreviation: NA, not applicable.
cross-sectional study showed considerable variation in power reactivation, but the frequency of use was not reported. Another powered mobility was commonly used for shopping and recreation seem to be the greatest risk related to the use of a motorized scooter. Training when providing a motorized device accounted for the majority of the wheelchair-related accidents for motorized scooters are extraordinarily limited. In a cross-sectional survey of older adults, Brandt et al found that powered mobility was commonly used for shopping and recreation, but the frequency of use was not reported. Another cross-sectional study showed considerable variation in power wheelchairs usage, even among reportedly full-time users. Izouni provided a qualitative view of the ability of powered mobility to restore independence and self-esteem. This study provides prospective, comparative data showing that provision of a motorized scooter to ambulatory persons with arthritis for use ad lib resulted in modest and selective usage. The scooters were predominantly used to access medically necessary locations. Over 50% of the scooter recipients in this study used the scooter to access food stores, and 30% to 50% used it to access doctor’s offices.

A significant concern in prescribing motorized scooters is the risk of harm. Scooter prescription, especially for ambulatory patients, is particularly controversial because of their high cost and concerns about deconditioning with scooter usage. This study was powered to detect a clinically significant short-term change in physical performance, and the scooters did not result in a statistically significant difference in physical performance. Moreover, 6MWD did not decline in physical performance in any of the subpopulations examined among the scooter users. However, minor collisions were relatively common. Nearly 20% of subjects in this study reported some kind of accident with the scooter, despite instruction in scooter use by experienced personnel. Accidents among powered mobility users in other studies include a 12% rate of “mishaps” found in a cross-sectional study among electric mobility users of public transportation, a 6% rate of tips/falls found with a test dummy on a standardized track, and powered mobility devices accounted for the majority of the wheelchair-related reports submitted to U.S. Food and Drug Administration from 1975 to 1993. Thus, accidents rather than deconditioning seem to be the greatest risk related to the use of a motorized scooter. Training when providing a motorized scooter may need to address specifically the potential for accidents and include practice using the scooter in crowded situations, over uneven ground, and loading and unloading the scooter.

One aspect of our findings that requires further consideration is why the scooters were not used more when the study subjects self-selected themselves as being interested in using a scooter in the first place. One possibility relates to limitations in the typical clinical practices for providing mobility aids in the United States. A recent study showed that instruction by expert personnel can increase the use of wheelchairs compared with usual care. Scooter usage and related benefits might have been greater had the study examined “best practice” rather than typical clinical practice with instruction being provided by the vendor rather than a clinician. On the other hand, the benefits of the scooter may be relatively limited for people who are able to ambulate, albeit with difficulty. Verbrugge and Svav used data from the National Health Interview Survey to empirically examine the relation of the use of equipment versus human help on perceived difficulty, fatigue, pain, and the amount of time required to perform various self-care tasks. They found that people who relied on equipment alone reported with greater difficulty, fatigue, pain, and time to complete mobility tasks than those who used human help. Thus, the price of self-sufficiency enabled by equipment may be fatigue, slowness, and pain, which potentially trade-off against the psychologic benefits gained from independence. Gignac et al examined perceptions of independence related to various coping strategies in over 200 persons with arthritis, including compensatory strategies such as equipment use, substituting help from another person, optimizing one’s physical abilities (eg, through exercise), or selecting which activities to perform and which ones to avoid. Compensation with equipment was the most common approach used for mobility, followed by optimization, then selection, and finally the use of human help. Compensatory strategies for mobility were moderately correlated with feelings of independence, dependence, and helplessness. In contrast, the use of human help and optimization strategies like exercise for mobility were not correlated with feelings of independence or dependence. Rosenfield and Faircloth reported similar findings in a qualitative study with 12 men and women with arthritis who described coping strategies for mobility as impacting their very identity. For the interviewees, movement was profoundly social in nature; it not only allowed for accomplishment of specific tasks, but it was an important

### Table 4: Descriptive Analyses of Scooter Use at Specific Locations Over Time

<table>
<thead>
<tr>
<th>Location</th>
<th>Scooter (n=24)</th>
<th>1 Month (n=23)</th>
<th>3 Months (n=22)</th>
<th>Control (n=23)</th>
<th>1 Month (n=22)</th>
<th>3 Months (n=21)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Food store</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Went to food store (%)</td>
<td>100.0</td>
<td>95.6</td>
<td>86.4</td>
<td>95.6</td>
<td>86.4</td>
<td>100.0</td>
</tr>
<tr>
<td>Used scooter at food store (%)</td>
<td>16.7</td>
<td>59.1</td>
<td>52.6</td>
<td>9.1</td>
<td>26.3</td>
<td>9.5</td>
</tr>
<tr>
<td><strong>Doctor</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Went to doctor (%)</td>
<td>66.6</td>
<td>65.2</td>
<td>63.6</td>
<td>43.4</td>
<td>72.7</td>
<td>71.4</td>
</tr>
<tr>
<td>Used scooter at doctor (%)</td>
<td>0.0</td>
<td>53.3</td>
<td>35.7</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td><strong>Friend/relative’s house</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Went to friend/relative’s house (%)</td>
<td>95.8</td>
<td>82.6</td>
<td>86.4</td>
<td>91.3</td>
<td>77.2</td>
<td>95.2</td>
</tr>
<tr>
<td>Used scooter at friend/relative’s house (%)</td>
<td>0.0</td>
<td>28.3</td>
<td>15.8</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td><strong>Church</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Went to church (%)</td>
<td>83.3</td>
<td>82.6</td>
<td>72.7</td>
<td>60.9</td>
<td>54.5</td>
<td>52.4</td>
</tr>
<tr>
<td>Used scooter at church (%)</td>
<td>0.0</td>
<td>31.4</td>
<td>18.7</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

*Percentage is calculated based on the number of subjects who went to that location.*
way to determine one’s value and commitment to health. Thus, the use of assistive technology may be closely tied to emotional perceptions about disability. In another qualitative study, Clemson et al. asked women who refused to implement environmental modifications to reduce their risk of falling why they decided not to follow therapist recommendations that might enhance their mobility. The women noted a number of factors as influencing their decisions, but the reasons all centered on perceptions of risk and beliefs about the ability to mediate these risks through behavioral changes. Treadwell and Lenert provided an interesting perspective, describing health decisions as being based on relative values rather than absolute values and that these relative values are nonlinear and affected by one’s baseline state. They proposed that the perceived benefits of wheelchair mobility to someone who is bedridden will be greater than with the perceived benefits of being able to walk for someone who is already using a wheelchair. Thus, for someone who is largely bed bound, the increased burdens of fatigue, pain, personal perceptions of dependency and disability, and even accidents, may be well worth-while given the severe preexisting limitations in independence. However, for someone who has even limited ambulation and is able to attain mobility goals with occasional help from another person, the personal trade-off may be less favorable.

Study Limitations

The study has a number of limitations. The study had adequate power to detect a clinically important change in physical performance with scooter usage over the short term, but it was not powered to detect minor changes in physical performance or changes that may occur over years of scooter usage. Data pertaining to the reliability and validity of the measures of scooter usage and mobility methods are limited. Other studies show that objective and self-reported measures of physical function correlate moderately well with one another among wheelchair users and among persons with physical disability. Despite randomization, there was a statistically significant difference at baseline in age and in premorbid wheelchair use, although all other variables showed no differences. However, there is no reason to believe that age or premorbid wheelchair use per se would interact with scooter usage in such a way as to affect the outcomes. This study does not speak directly to the potential impact of scooter usage on specific body systems. The 6MWD correlates well with measures of balance, muscle strength, and endurance. Study results may not be generalizable to people with more severe disability or with other medical conditions. The use of 3-month recall for accidents using the scooter may be biased because of inaccurate recall. Training in scooter usage did not take place under controlled circumstances because of the study goal of examining “effectiveness” of scooters rather than efficacy. Enhanced training might reduce the high accident rates observed in this study. The study took place over a prolonged period of time, during which scooters became increasingly available in commercial venues, which may have diluted the overall effect of the intervention. Finally, the optimal population in which to examine deconditioning with use of a motorized scooter (ie, ambulatory persons who are not deconditioned) may not be the best population in which to determine the benefits from the device (ie, persons with severe mobility limitations).

CONCLUSIONS

This is the first study to use rigorous methodology to provide prospective data on real-life usage of powered mobility, along with attendant benefits and risks. The study shows that the activities commonly accessed with a scooter by ambulatory people with arthritis appear to be a mixed group that includes activities related both to personal convenience (eg, going to the mall) and to medical necessity (eg, doctor visits). In this ambulatory population, the impact of provision of a scooter on physical performance and on overall levels of activity appeared to be no more than modest. Although the scooter users used their scooters nearly every day, they were not used for such a prolonged period of time as to have a clinically significant effect on physical performance over the study period. Thus, the scooters may function predominantly to facilitate transitions between activities, enabling access to a broader selection within the same spectrum of activities, whereas the activities themselves are performed in a standard fashion. Alternatively, for people with minimal to moderate mobility impairment, the price of independence with a scooter may not trade-off favorably against the greater ease of having some help with tasks that have a strong social component and ready availability of social support (eg, visiting friends).

Nationally, rates of disability have declined at the same time as rates of equipment usage have increased for those same activities, suggesting that priority be placed on research focused on assistive technology and physical environment. Empirical data are needed on the effects of diverse approaches to help with mobility, from assistive devices, to elder housing built on principals of universal design, to city planning for enhanced walkability. Data are needed across the broad spectrum of potential users of such interventions, ranging from the most disabled to the least disabled. The effect of these interventions should be examined not only in the laboratory but also in the context of the daily lives of people with diverse physical abilities. Future studies might fruitfully examine methods for training patients in the safe use of scooters and lifts, identifying locations in which scooters may or may not be safely used and identifying clinical factors that may modify recommendations for safe scooter use. With such data in hand, we will be able to better inform persons with arthritis and other conditions as they make personal decisions about coping with mobility impairment.

Acknowledgments: We thank Michael Zolkewitz and Katina Hargraves for assistance with data collection and data management and William Logan for editorial review.

References


Suppliers
a. Pride Mobility Products Corp, 182 Susquehanna Ave, Exeter, PA 18643.
b. Bruno Independent Living Aids Inc, PO Box 84, Oconomowoc, WI 53066.
c. Model HJ102; Omron Healthcare Inc, 1200 Lakeside Dr, Bannockburn, IL 60015.
Spasticity Experience Domains in Persons With Spinal Cord Injury

Jane S. Mahoney, DSN, RN, Joan C. Engebretson, DrPH, RN, Karen F. Cook, PhD, Karen A. Hart, PhD, Susan Robinson-Whelen, PhD, Arthur M. Sherwood, PhD


Objective: To understand the everyday life experiences of persons who have spasticity associated with spinal cord injury (SCI).

Design: Applied ethnographic design.

Setting: Patients’ homes and rehabilitation clinics.

Participants: Twenty-four people with SCI who experience spasticity.

Interventions: Not applicable.

Main Outcome Measures: Domains identified through qualitative analysis of in-depth open-ended interviews.

Results: Domain analysis revealed 7 domains: physical, activity, emotional, economic, interpersonal, management, and cognitive. Descriptive subcategories within each domain were identified. Patients personalized the meaning of spasticity and expressed their understandings of the condition in ways that may not be consistent with clinical definitions. Some patients suggested that being able to control spasticity was preferable to total suppression.

Conclusions: Spasticity-related interventions need to be aimed at what matters most to the patient. It is critical for clinicians to understand patients’ experiences to make accurate assessments, effectively evaluate treatment interventions, and select appropriate management strategies. When providers reconfigure patients’ descriptions to fit neatly with a biomedical framework, they may not be consistent with clinical definitions. Some patients suggested that being able to control spasticity was preferable to total suppression.

Key Words: Muscle spasticity; Rehabilitation; Spasm; Spinal cord injuries.

© 2007 by the American Congress of Rehabilitation Medicine and the American Academy of Physical Medicine and Rehabilitation

Spinal Cord Injury (SCI) is a devastating condition with life-long consequences. Motor control problems related to SCI are typically labeled spasticity. In surveys related to perceptions of problems associated with SCI, patients consistently ranked spasticity among their top 4 life concerns.

Because spasticity is a major issue for patients after SCI, assessment of its impact is worthy of concentrated, disciplined attention.

There is no criterion standard for assessing the severity of spasticity or consensus on its definition. The most-often cited definition was offered by Lance who described spasticity as velocity-dependent increased resistance to passive stretch as a component of upper motoneuron dysfunction. More recently, spasticity has been described as disordered sensorimotor control as a result of an upper motoneuron lesion. The condition presents as involuntary movement that may be constant or intermittent.

Neither of these descriptions addresses the impact of spasticity on the life of a person.

The fact that commonly used clinical scales of spasticity do not correlate well with each other highlights the lack of conceptual clarity regarding spasticity. A condition must be clearly defined before it can be precisely and reliably measured. In this regard, we distinguish between evaluations of spasticity that document its manifestations and severity and those that measure its impact on persons’ daily lives. It is the former that has preoccupied investigators up until now and the latter that needs, and deserves, scholarly attention. The manner in which spasticity alters the life of a person with SCI should provide the basis for clarifying spasticity’s definition and improving its measurement. This is consistent with recommendations of Burridge et al who recommended including the impact of spasticity on activities of everyday living in spasticity measurement.

This study was part of a larger project aimed at developing an instrument to measure the impact of spasticity in persons with SCI. The purpose of this study was to identify domains related to the everyday life experiences of persons who have spasticity associated with SCI (1) to inform item construction and (2) to provide a description of aspects of the experience to clinicians and researchers. This approach incorporated the voices that inform what matters most to the consumer or patient stakeholder. The analysis reported in this study translates the insiders’ views of living with spasticity into domains. These domains can serve as the foundation for clinical assessment and instrument development. Familiarity with these domains may broaden clinicians’ frames of reference in caring for persons who experience spasticity and increase awareness of issues that may influence the effectiveness of treatment.

Methods

Design

For this study, we used an applied ethnographic design, a form of qualitative research. Ethnography is widely used to understand the experiences of living with a health condition.
Such a design allows for exploration of an experience from the perspective of the one who lives with the phenomenon.8,9 The underlying assumption is that the insider, not the outsider, is the expert about what it is like to live with the condition. The term insiders refers to those who live with a condition as opposed to outsiders who look in on someone else’s experience.7,8,13

**Sampling**

A purposeful sampling technique was used. Purposeful sampling is aimed at the selection of information-rich cases that will elucidate the phenomenon under investigation.14 This method has been used in other ethnographic studies.10,15 In this study, the purposeful sampling technique of criterion sampling14 was used. Subjects meeting the criteria, which included English-speaking adults, with a diagnosis of SCI and self-reported spasticity who were at least 1 year postinjury, were purposefully selected.

**Procedure**

After receiving approval from the institutional review board, participants from previous research studies were recruited. These studies, conducted at Baylor College of Medicine (Houston, TX) and the Michael E. DeBakey Veterans Affairs Medical Center (MEDVAMC, Houston, TX), evaluated the long-term psychosocial and physical effects of SCI, including the experience of persons with spasticity after SCI. Additional participants were recruited from a list of patients recently receiving care through the MEDVAMC SCI Service Line. Thirty-five potential participants were contacted by telephone, informed of the study, and invited to participate in the study. Of those, 8 declined because of lack of interest or bad health, and 2 could not be reached subsequently for scheduling despite repeated efforts. The remaining 25 were interviewed; however, the tape recorder failed during 1 interview. The decision was made to eliminate this subject because of the lack of reliable data. Thus, the sample size used in this study was 24 (68.5% of potential sample). Participants selected a time and location that was private and convenient for them. The locations included informants’ homes and private areas in outpatient clinics. Participants were compensated $25 for their time.

An investigator experienced in qualitative interviewing techniques (KAH) administered a short, demographic questionnaire and conducted open-ended semistructured in-depth interviews lasting approximately 45 to 90 minutes. Interview questions were designed to elicit persons’ experiences with spasticity in everyday life. Based on recommendations by Spradley,8 the interview guide consisted of questions and probes designed to move from a general overview of the experience, “grand-tour” questions, to more focused questions including probes for describing, clarifying, contrasting, and verifying. Examples of questions are presented in Table 1. Interviews continued until saturation and redundancy of the data were reached. Data were determined to be redundant when no new information was obtained. Saturation was reached when multiple variations in accounts were identified, critiqued, and explained.16

Consistent with the emergent design inherent in qualitative research, interview questions were added as new data were collected. For example, in initial interviews, when respondents were asked about spasticity, they often referred to spasticity as spasms and involuntary movements. Once we discovered that such terms used were part of the vernacular of the “insider,” the interview protocol was modified to include inquiries about spasms and other involuntary movements.

The interviewing investigator also constructed field notes consisting of important observations not noted in the transcriptions, such as details of the physical settings, nonverbal cues, and theoretical notations. Interviews were audiotaped and transcribed by a trained transcriptionist experienced in medical and qualitative interview transcriptions. An independent analyst (JSM) compared the transcripts and audiotapes to ensure authenticity.

**Analysis**

In keeping with the overall goal of the study to guide the development of an instrument as well as to describe the experience, a domain analysis, a structured analytic approach described by Spradley,8 was used to identify important domains and categories, one of the main purposes of in-depth ethnographic interviewing.8,17 This method contributes to understanding the shared problems and problem-solving strategies of a particular group of people,3 such as those who experience spasticity. This method can be used to recognize factors for instrument development.17

Two analysts experienced with qualitative research methods (JSM, JCE) examined the data independently. Transcripts and field notes were read and reread. Codes were written in the margin of the text. A list of codes was examined for relatedness. The textual data were then organized into categories based on related codes. Next, we constructed tables for each category for further examination. In this way, we created an audit trail that allowed for a stepwise approach to refinement. The refined categories were scrutinized for semantic relations and organized into domains.8,17,18 Throughout the process, the analysts engaged in analytic discourse until agreement was met, resulting in a consensus of the findings.

Qualitative research is concerned with the trustworthiness and credibility of the data and analysis, concerns that are comparable to issues of reliability and validity intrinsic to quantitative research.19 For the current study, validation of findings included a craftsmanship approach that incorporated continually checking, questioning, and theoretically interpreting findings, strategies that have been suggested as a means of quality control.20 In this study, trustworthiness was also enhanced through verification of the findings with the respondents, triangulation of data between and among narratives, and ongoing peer debriefing among members of the study team.

---

**Table 1: Sample Interview Questions**

<table>
<thead>
<tr>
<th>Scope</th>
<th>Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grand tour</td>
<td>Please tell me about your experiences with spasticity.</td>
</tr>
<tr>
<td>Focused</td>
<td>Can you describe a particular time you experienced spasticity?</td>
</tr>
<tr>
<td></td>
<td>What happens when you are around people and you experience spasticity or think you are about to experience it?</td>
</tr>
<tr>
<td></td>
<td>Can you describe what you think is happening in your body when you are having spasticity?</td>
</tr>
<tr>
<td></td>
<td>What do you think causes the involuntary movements?</td>
</tr>
<tr>
<td></td>
<td>Are there things you need to have to manage spasticity?</td>
</tr>
<tr>
<td></td>
<td>How did you learn about these things?</td>
</tr>
<tr>
<td></td>
<td>Can you give me an example of a bad time with spasticity?</td>
</tr>
<tr>
<td></td>
<td>Can you give me an example of a “not so bad” time?</td>
</tr>
</tbody>
</table>

---

of the issues as fluid and context dependent. In fact, central to participants’ descriptions was the experience these domains as discrete entities. This was not at all the case. As may suggest, inappropriately, that participants experienced the totality of the issues raised, the gathering of the domains into a list entwined. Although the following listing is an accurate inventory of the relationships among domains were often overlapping and sometimes little overlap.

TABLE 2: Summary of Sample Characteristics (N = 24)

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y)</td>
<td>45.1 (21–68)</td>
</tr>
<tr>
<td>Years of education</td>
<td>13.0 (8–18)</td>
</tr>
<tr>
<td>Years postinjury</td>
<td>16.0 (1–39)</td>
</tr>
<tr>
<td>Level of injury</td>
<td></td>
</tr>
<tr>
<td>Paraplegia</td>
<td>11</td>
</tr>
<tr>
<td>Tetraplegia</td>
<td>13</td>
</tr>
<tr>
<td>Low</td>
<td>7</td>
</tr>
<tr>
<td>High</td>
<td>6</td>
</tr>
<tr>
<td>Race</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>13</td>
</tr>
<tr>
<td>Black</td>
<td>7</td>
</tr>
<tr>
<td>Hispanic</td>
<td>3</td>
</tr>
<tr>
<td>Asian</td>
<td>1</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>17</td>
</tr>
<tr>
<td>Female</td>
<td>7</td>
</tr>
<tr>
<td>Veteran status</td>
<td></td>
</tr>
<tr>
<td>Veterans</td>
<td>13</td>
</tr>
<tr>
<td>Nonveterans</td>
<td>11</td>
</tr>
</tbody>
</table>

NOTE. Values are mean (range) or n.

RESULTS

Seventeen men and 7 women between the ages of 21 and 68 participated. The demographic and clinical characteristics of the sample are summarized in Table 2. Although our initial investigation targeted spasticity, we became aware early in data collection that the reports of people who live with the condition extended beyond what is usually encompassed in clinical definitions of spasticity. Participants often described their experiences of spasms and involuntary movements that may or may not be included in definitions of spasticity. Data analysis revealed 7 domains associated with the experiences: physical domain, activity domain, emotional domain, economic domain, interpersonal domain, management domain, and cognitive domain. The ways respondents spoke about their experiences were sometimes different from and even contradictory to the biomedical understanding of the condition. For example, when asked to discuss spasticity, some informants included descriptions of pain, which, in some instances, may not fit the biomedical description of spasticity. In describing their view of the experiences, respondents blended folk and medical terms. Our findings suggest that patients tend to bundle symptoms associated with SCI and may not express these symptoms in language that is consistent with biomedical understandings of spasticity. The glossary of the patient living with the reality of the condition and that of the health care provider may have little overlap.

Interrelatedness of Domains

In reviewing the interview data, we were able to identify specific domains that described participants’ experiences. Although specific domains were identified, the patterns within and among domains were often overlapping and sometimes entwined. Although the following listing is an accurate inventory of the issues raised, the gathering of the domains into a list may suggest, inappropriately, that participants experienced these domains as discrete entities. This was not at all the case. In fact, central to participants’ descriptions was the experience of the issues as fluid and context dependent.

Physical Domain

The physical nature of involuntary movements was illuminated richly in the respondents’ accounts. The respondents spoke about the physical domain more often and in much more detail than other areas with particular emphasis given to describing the characteristics of the movements. Subcategories of the domain included the following: characteristics of movements; severity, pain, and other sensations; timing; course; and triggers.

Characteristics of movements. Respondents described several characteristics of involuntary movements (appendix 1). One man said, “I differentiate between tone which is stiffness, clonus which shakes if I open fingers or push on my foot, and spasms that are sudden shooting movements.” Another man blended the terms spasms and spasticity. “I have two or three different types of spasms . . . [one] that is constant . . . there seems to be a spasticity of movement . . . that is relatively constant in trying to reach forward. And then I have a spasm that is of a clonic type nature. It will shake, rattle, and roll. And then I have a spasm that is just a sudden, brief, out of the blue sort of electric shock type spasm in that it is a pure jolt. Or it may be that I am making a purposeful movement, and, suddenly, I will lose the ability to make that move. It is like holding on to a pen or pencil and suddenly dropping it or making a movement with my arm, and, suddenly, my arm will just drop away.”

Another man explained, “When I exercise too much, or when I am tired, especially in my hands, I get the real repetitious type of spasm, and I am unable to do things.” Another man said, “I believe you call them extorsory spasms, my arms will shoot out, and my hands will flop out to the side and shake. Sometimes, it will fall off my armrests, because it hits my stomach. It hits my diaphragm, so breathing is kind of uncomfortable . . . .” These contextual descriptions of characteristics of spasms included allusions to their impact on everyday living, for example, making breathing uncomfortable.

Severity, pain, and other sensations. Several respondents showed their difficulty in accurately articulating the sensation. One participant stated, “It’s a feeling you can’t pinpoint.” Another called it, “Indescribable. If you’re not in this shape you can’t imagine what all the turmoil is.” Comments like these suggest that for people who have experienced an altered state of embodiment, such as SCI, language is often inadequate to explain sensations.

Some used metaphors to describe somewhat vague sensations: “seems like something tied around my waist,” “like ants crawling on you,” and “like something gnawing in your leg.” Some of these descriptions may be interpreted by clinicians and researchers in some instances as sensations unrelated to spasticity. However, people living with the condition do not adhere to biomedical categorizations of sensations but rather attempt to describe sensations as they experience and interpret them. The severity of spasms and spasticity was described in terms of pain and other sensations. Some experienced no sensation, whereas others described sensations that were mild, moderate, or severe (appendix 2). Several described violent expressions of their spasms and/or spasticity. One participant described symptoms “severe enough that they will knock my breath out and shout my arms out. I mean, unpredictable.” Another described a similar experience, “My spasms have thrown me out of my chair on numerous occasions. Where I have broke both of my legs, and I have gotten thrown out of my chair to where . . . the whole chair turned over.”

The apparent severity of the symptoms was distinguishable from the severity of their impact. One person described...
the impact of spasms and spasticity as “more or less bothersome . . . something that comes with the territory.” The same person described a dramatic experience: “The spasms, sometimes they get so bad that I have a little short repeating type of spasm or a long kind that will take my body almost and throw me out of the chair and more or less, I call them an obstacle . . . a challenge . . . whatever.” These excerpts from the same respondent show the dependency of the impact of experience on the person’s individual context. Spasms strong enough to “nearly throw [one] out of the chair” were described as merely “bothersome.” Perhaps relative to his other experiences, being thrown out of a chair for this participant is merely “a challenge.”

Many participants described positive impacts of spasms and spasticity including prevention of muscle atrophy. Some appreciated having a sense that something was “going on” in the affected region. Severe spasms served as an alert to some participants of the possibility of a urinary tract infection (UTI) or some other problem. Others noted the assistance spasticity provided in making transfers. Because of the desirable impacts of spasticity, several stated that they would not want to be rid of their spasms and spasticity. They would prefer having the ability to initiate the spasms and spasticity at the times when they would be helpful. This paradoxical response to spasticity as being simultaneously hurtful and helpful is especially noteworthy and further shows the complex nature of the experiences.

**Timing.** Timing, frequency, and duration of spasms and spasticity varied and were context dependent. For some, the occurrence was rare, whereas for others it occurred quite often, “sometimes 30 to 40 times a day.” The spasms might last only a few seconds at a time or result in a feeling of “constant tone.” Some experienced more spasms in response to specific health conditions, and being touched by others. Transfers often stimulated involuntary movements, “When you first lay down from sitting are the worst times.” This was confirmed by others who identified moving to a prone position as a factor. One man said, “If I’ve been up a long time, I get more spastic.” Some talked about increased involuntary movement after therapy or after lifting weights, “when your body gets pumped up.” Others said that the spasms were worse when they did not exercise.

Weather changes, rainy weather, cold weather, and cool temperatures from air conditioning were reported by some as contributing to worse experiences, but others’ symptoms worsened when the air temperature was warm. Additional exacerbating factors included rolling over a bumpy surface such as a sidewalk and wearing tight-fitting clothes or shoes. One participant reported, “Sometimes my skin is so sensitive, just by pulling the sheet on to cover my knees causes spasms.”

Touch from others was identified as a cause of spasms and spasticity, particularly when the touch was unexpected. Others talked about worsened symptoms when they were tickled, held in a particular way, hugged, or when their feet were touched. One participant reported, “When someone touches me—if it is a hard touch it will spasm, but not always with a soft touch.”

**Activity Domain**

Reports varied with respect to the effect involuntary movements had on activities. Some reported that they “just tolerate them and go on.” Others made substantial modifications including not planning activities when the symptoms were severe. One respondent reported the impact on driving: “Sometimes driving, I just have to pull over and sit a while and relax if my muscles are having a lot of spasms.” Another reported, “On that day when the muscles are bad, rainy days or when the weather changes. . . . It is worse on those days, and then sometimes on those days, I don’t attempt to do anything, because I have an increase in spasms.” Another explained that involuntary movements “can make me go back and fall over, ‘cause I have no balance. . . . I lose complete control and can’t do anything until it is over.”

**Emotional Domain**

The emotional experiences also spanned a continuum. Descriptions ranged from “nothing to fear or worry about” to “don’t feel like doing anything.” Specific emotions described by the respondents included embarrassment, frustration, uncertainty, worry, moodiness, crankiness, and powerlessness. One participant described the complexity of emotions brought on by the involuntary movements: “At first, it is scary because you don’t know . . . the spasm may come up in a portion of your body that is going to make you fall . . . tension and nervousness . . . in fact, I do have a lot of paranoia about it. This is more or less a mental thing . . . it takes so much not to really think about it.”

**Economic Domain**

The economic impact of involuntary movement was experienced both in decreases in earning potential and increases in the cost of managing their condition. Some described job interviews in which they were embarrassed by spasms that they felt influenced a negative outcome. Others said that the movements interfered with particular job tasks, limiting their work options beyond the impact of SCI alone.

Some respondents reported that they had adequate medical insurance, but several indicated a need for better coverage of medicines and physical therapy. Several indicated a need for reimbursement for complementary and alternative therapies (e.g., massage therapy).

**Interpersonal Domain**

Respondents described the interpersonal impact of spasms and spasticity according to (1) people they needed to help them with the movements, (2) what people thought about the
movements, and (3) interruptions in interpersonal encounters (appendix 3). One man reported needing more assistance because his involuntary movements created a lack of confidence in performing necessary movements. Another man stated, “I may hit a bump in the sidewalk or a little dip in the road, it may send off a spasm, and my arm may flop off to the side, or I may lean over to the side and I am kind of stuck until someone can come by and help put my arm back or push me over.” Others poignantly described how the movements interfered with pleasurable activities spent with loved ones. Additional remarks highlighted the lack of public understanding about this condition and suggested an experience of being socially stigmatized, “A lot of people don’t like to touch me.” Remarks such as these represent a shift from the more typical medical discourse because the respondents use a powerful other language to highlight the impact of spasticity on interpersonal aspects of everyday life.

Management Domain

Issues related to functional self-management included the following: (1) need for unique self-care techniques, (2) need for greater assistance, and (3) incorporation of medical and alternative therapies. Often participants reported adapting medical management strategies to make them fit within their individual self-management regimen. These strategies were person and situation specific and included strategies for modulating, preventing, and stopping involuntary movements (table 3), as well as initiating them when their effects were needed (eg, during transfers).

Although some emphasized the benefits of adhering to prescribed medication regimens, others described the need to modify them. One participant reported, “The doctor agreed to let me regulate my medication regimen because I know more about myself than they could ever know.” Some talked about only taking their medications when they were “going out” to reduce the chances of having a spasm away from home. Many respondents voiced concerns about becoming addicted or dependent on some of the medicines, disturbed by the sedating side effects associated with many of the drugs and the feeling that they were “doped up.”

Many reported using alternative therapies such as massage, acupressure, acupuncture, deep breathing, meditation, sauna, and whirlpool. Several discussed using marijuana as an effective strategy. “It makes you relax and decreases the intensity of them.” “Pharmaceuticals can’t compare with Mother Nature; I know that for a fact. I can feel it in my heart. Pot did not keep me in the house or too mellow to do anything like Valium.” However, many were concerned about the legal implications associated with using marijuana. “I would love to use it but wouldn’t risk having a blood test positive for pot at the VA and consequences such as not getting a federal loan for college if it is on your record.”

Several respondents described spiritual or religious approaches for dealing with their spasms and spasticity. Some talked about listening to gospel music or watching religious television shows as a way of coping with their condition and with their spasms and spasticity. One participant reported coping by getting “closer to God.” Another said, “There is a reason for everything, and I give it to God. I don’t complain to Him, I just tell Him. I talk to Him just like I am talking to you, and that is about it.”

Cognitive Domain

Informants described their interpretations of what happens when they experience involuntary movements. Within these descriptions is the meaning they ascribe to the movements.

<table>
<thead>
<tr>
<th>Management</th>
<th>Manipulations and Modifications</th>
<th>Medical Strategies</th>
<th>Alternative Therapy Strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ways to prevent</td>
<td>“If I know when I am going to have a spasm I stay still, and I won’t spasm.”</td>
<td>“Baclofen pump.”</td>
<td>“Marijuana.”</td>
</tr>
<tr>
<td></td>
<td>“Range of motion.”</td>
<td>“Medication.”</td>
<td>“Relaxation.”</td>
</tr>
<tr>
<td></td>
<td>“A regular good bowel program.”</td>
<td>“Surgery, but I don’t want to have another surgery.”</td>
<td>“Acupuncture,” “Acupressure.”</td>
</tr>
<tr>
<td>Ways to modulate</td>
<td>“Stretching”</td>
<td>“Physical therapy”</td>
<td>“Pray.” “Leave it up to the Lord.”</td>
</tr>
<tr>
<td></td>
<td>“Bend over and put my head between my legs.”</td>
<td>“Physical therapy”</td>
<td>“Massage.”</td>
</tr>
<tr>
<td></td>
<td>“Change position.”</td>
<td>“Medication.”</td>
<td>“Meditate.”</td>
</tr>
<tr>
<td></td>
<td>“Breathe. A lot of people tense up, but breathing helps mellow them out.”</td>
<td>“Exercise.”</td>
<td>“Acupuncture/acupressure.”</td>
</tr>
<tr>
<td></td>
<td>“Exercise. Do my range of motion.”</td>
<td>“Marijuana.”</td>
<td>“Marijuana knocks them out instantly . . . I don’t have to take meds if I smoke.”</td>
</tr>
<tr>
<td>Ways to stop</td>
<td>“Raise the bed a little bit.”</td>
<td>“Medication is the only thing,” baclofen (oral and pump), Neurontin.</td>
<td>“Valium. I’m not the same person, but they work.”</td>
</tr>
<tr>
<td></td>
<td>“Keep exercising. They will stop.”</td>
<td>Valium.</td>
<td>“Alcohol.”</td>
</tr>
<tr>
<td></td>
<td>“Change position.”</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>“Lift my toe.”</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>“Push on it.” “Someone else will push on it and make it stop.”</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>“Take a deep breath and body spasms stop.”</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>“Heat” (whirlpool, heating pad, and so on).</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Some described them as a way an injured part of the body communicates with the brain. “When I have a spasm, it excites the nerves and sends pain messages to the brain—‘Nerve Damage’.” Another person explained, “The nerve reaction doesn’t get up to your brain, so it doesn’t know how to react. It just causes a cycle, round and round until it stops.” Along the same lines, another man explained, “My body is trying to send a signal to my brain. I don’t think it gets past the breaking point in my spinal column, so maybe it goes back down through my body, and it goes back up, and it can’t go past that breaking point, so it is kind of reverberating [sic] back and forth.”

Others suggested the movements were a reaction to something unpleasant. For example, “The body is not liking what it feels. I can’t feel hot or cold on my legs. When they put the cold water...my body doesn’t like it, so I will spasm.”

As previously stated, some liked the feeling that something was “going on” in the injured parts of their bodies. For example, one person described involuntary movements as “just my body telling me I am still alive.” Others described similar interpretations: “My body is trying to get a little movement,” “It tells me my arm is tired,” “My body is telling me something is wrong.”

The variations in interpretations show the uniqueness of the experiences to individual participants. Although some people found the experience of living with spasms and spasticity unbearable, others minimized it: “An occasional spasm is now a part of who I am.” “It’s just one of those things you accept.” Others accepted that there were both advantages and disadvantages in the experience.

Issues of control emerged as central to the way participants understood spasms and spasticity. One respondent noted that, “It gets worse if you sit back and let it take control,” suggesting that control and a sense of self-efficacy are related to the person’s experience of living with these movements. “You’ve got to control it—fight back and live a normal life.” The idea of “fighting” the spasms was frequently used to describe the intensity of the emotional and physical work required to manage the movements. Some patients suggested that being able to control the movements was preferable to eliminating them altogether, such as results through the use of currently prescribed pharmacologic and surgical interventions.

**DISCUSSION**

Spasticity is a multifaceted phenomenon that has been a challenge to capture and adequately measure. In following the recommendations of Brown and Gordon,7 we recognized the importance of incorporating the insider’s voice in instrument development and clinical practice. The domains identified in this study are useful for instrument development as participants articulated important information about the impact of spasticity in everyday life. The domains allow for an understanding of the experiences beyond physical measures of muscle activity. Descriptions of physical, activity, emotional, economic, interpersonal, management, and cognitive domains present a holistic image of what people experience as they live with spasticity associated with SCI.

Participants personalized the meaning of spasms and spasticity and expressed their understandings of it in ways that may or may not be consistent with clinical definitions of spasticity. Our findings highlight everyday language used to express how people interpret spasticity in everyday life. People with SCI, as in this study, often do not always express spasticity symptoms in language that is consistent with the biomedical terminology and understanding. For example, even though spasms are not a component of the biomedical definition of spasticity, the term *spasm* was often used interchangeably with spasticity in participants’ accounts.

This supports the view that the perspective of the insider may vary from that of the outsider.7 The existence of such incongruencies between patients’ (insiders) and providers’ (outsiders) understandings are consistent with Kleinman’s framework21 for explanatory models. Explanatory models are persons’ ways of interpreting and explaining a phenomenon. Patient taxonomies of symptoms are often different from those of the clinician. Therefore, clinicians often need to translate the patient’s experience into the nomenclature of clinical understanding. Failing to attend to this translation step can lead immediately to failed communication and, eventually, to suboptimal treatment. It is critical that portions of the patient’s account that do not fit neatly into the biomedical explanation are not discounted by the clinician or the researcher. Clinicians who are solely oriented to the medically authorized narrative often consider these patient accounts irrelevant; however, legitimizing the patient’s experiences is a key task in the provision of care.22

Some participants used medical language to a more extent than others. This may be explained by the fact that people who have lived with a damaged spinal cord for many years have been influenced by multiple health care providers adapting portions of their clinicians’ glossary but also modifying it to make it their own. Furthermore, perhaps because of the unpredictable nature of the precipitating event and consequent lack of uniformity in the nature of their lesions, there is substantial variability in how individuals experience spasticity. This variability confounds attempts to define and assess spasticity. Although recent contributions by the European SPASM group4 offer some hope in this latter effort, they did not fully avail themselves of the benefits that may be obtained through use of the International Classification of Functioning, Disability and Health,23 which appropriately focuses attention on the separate domains of structure and function (where attention is most often given for spasticity), activities, and participation, defined as involvement in a life situation.6 The findings reported in this study offer insights that may be useful in expanding the understanding of the life situation associated with spasticity in persons with SCI. Additionally, findings from this study will be used to guide instrument development aimed at measuring the impact of spasticity in everyday life. Focus groups are planned to refine the domains, and psychometric testing will follow.

Although we were able to use participants’ reports to identify domains associated with the person’s experience of spasms and spasticity, it is extremely important for the dynamic and integrative nature of these experiences to be recognized. This is particularly important when attempting to appropriate these translations of participants’ experiences into a biomedical understanding of spasticity in planning patient-centered interventions. Participants who have reported a preference for control rather than total suppression of spasticity suggest an alternative approach to the treatment of spasticity that highlights the possibility of some form of restored control over one’s body and consequently improving the impact of spasticity to everyday life. This finding has direct clinical implications because clinicians should explore patients’ preferences for treatment and consider the impact of living with spasticity in everyday life as a complex experience that requires attention be given to the domains in the personal experience that matter most to the patient. This line of thinking creates a shift from the “fix-it” model associated with acute conditions to one aimed at indi-
visualizing interventions focused on control rather than suppression, a view often associated with chronicity. Such interventions would have the added value of improving quality of life and addressing patient preferences.

CONCLUSIONS

Participant’s descriptions of their experiences with spasticity provided insight into its nature and the problems associated with it. These experiences can be understood according to 7 domains: physical, activity, emotional, economic, interpersonal, management, and cognitive. Informants personalized the meaning of spasticity and expressed their understandings of the condition in ways that may or may not be consistent with clinical definitions. They articulated benefits of spasticity that are not often acknowledged in the SCI and plasticity literature.

APPENDIX 1: CHARACTERISTICS OF MOVEMENTS

**Strong sudden movement**

- “I am sitting up one minute, and then all of a sudden, I am flopped over.”
- “My arms and legs will shoot out straight.”
- “My knee will come up and my leg will shoot back out.”
- “It’s like a pure jolt.”
- “There is sudden tightening.”
- “It tries to throw me out of my wheelchair. It throws me forward.”
- “It can fold me over like a sandwich.”

**Repetitious**

- “I get a short repetitious type.”
- “You will feel tightness in your body and your legs will either kick or shake.”
- “Mine are real minor. My arms just shake a little bit.”
- “My legs shake a little bit.”
- “My legs start to move, tremble, jump around like popcorn popping.”
- “My knees start twitching.”

**Constant tone**

- “Tone keeps me stiff, and my arm slowly falls to the side.”
- “I have more a muscle tone where my legs will stiffen up.”
- “Spasms come and go, but the tone is constant.”

**Interferes with deliberate movement**

- “My arms will fall immediately so I can’t move fluidly.”
- “The leg is restricted so I can’t move it any farther.”
- “They can make you move in a way you don’t want. Makes you do things you don’t want to do.”
- “It doesn’t want to be moved.”
- “My leg moves by itself.”

**Involves certain body parts**

- “My stomach gets real tense, and it kind of comes up like you would think of bread rising, and then it will be real slow going back down until my stomach and back are relaxed again.”
- “The stomach squeezes from the inside out. It makes your whole body spasm from the inside out.”
- “There’s the total body type where everything freezes up.”
- “Spasm happens somewhere and it sends it through the back of my shoulders.”
- “It goes up my spine and down both my arms and up the back of my head.”
- “They are dominant in the weaker portion of the body.”
- “It goes from the bottom of the toes to the hip.”
- “It moves from the legs to the stomach.”
- “It happens in the legs, foot, knee, and ankle.”
- “Usually they happen in the trunk.”

**APPENDIX 2: SEVERITY, SENSATION, AND PAIN**

<table>
<thead>
<tr>
<th>Severity</th>
<th>Sensation</th>
<th>Pain</th>
</tr>
</thead>
<tbody>
<tr>
<td>No sensation</td>
<td>“I don’t answer the phone on bad spasm days.”</td>
<td>“I can’t go to the mailbox or the movies (with someone) because it is cold, and that will cause them.”</td>
</tr>
<tr>
<td>Mild</td>
<td>“Bothersome.”</td>
<td>“They probably think I am having a seizure.”</td>
</tr>
<tr>
<td>Moderate</td>
<td>“A little painful.”</td>
<td>“I need to reposition me.”</td>
</tr>
<tr>
<td>Severe</td>
<td>“Most suicidal . . . pain all the time . . . I’ve been through too much hell . . .”</td>
<td>“They hold my hand during a spasm, especially in public. It’s really cool.”</td>
</tr>
</tbody>
</table>

**APPENDIX 3: IMPACT TO INTERPERSONAL RELATIONSHIPS**

<table>
<thead>
<tr>
<th>Who Is Needed to Help</th>
<th>What Others Think</th>
<th>Interruptions in Interpersonal Encounters</th>
</tr>
</thead>
<tbody>
<tr>
<td>“I need someone to drive me when I am having spasms.”</td>
<td>“They probably think I am having a seizure.”</td>
<td>“I can’t go to the mailbox or the movies (with someone) because it is cold, and that will cause them.”</td>
</tr>
<tr>
<td>“I need my wife to reposition me.”</td>
<td>“Family and friends aren’t bothered.”</td>
<td>“I don’t answer the phone on bad spasm days.”</td>
</tr>
<tr>
<td>“She holds my hand during a spasm, especially in public. It’s really cool.”</td>
<td>“Strangers think it’s weird.”</td>
<td>“I apologize to the dentist. She had to get used to working on a moving target.”</td>
</tr>
<tr>
<td>“My father will come over and help with my spasms. He is getting older and has to rely on friends.”</td>
<td>“Think my legs are moving and I’m going to walk.”</td>
<td>“I can’t go out and talk to people, and I like to go out.”</td>
</tr>
</tbody>
</table>
APPENDIX 3: IMPACT TO INTERPERSONAL RELATIONSHIPS (cont’d)

<table>
<thead>
<tr>
<th>Who Is Needed to Help</th>
<th>What Others Think</th>
<th>Interruptions in Interpersonal Encounters</th>
</tr>
</thead>
<tbody>
<tr>
<td>“I need someone to come along and put my arm back.”</td>
<td>“They think you are making excuses.”</td>
<td>“In the middle of a conversation it hits. I will stop and focus.”</td>
</tr>
<tr>
<td>“God helps me.”</td>
<td>“A lot of people don’t like to touch me.”</td>
<td>“When a friend gives me a hug (I can spasm).”</td>
</tr>
</tbody>
</table>

References
Motor Points for the Neuromuscular Blockade of the Subscapularis Muscle

Tim P. Harrison, MBBS, BSc, Anna Sadnicka, MBChB, BSc, Deborah M. Eastwood, MB, FRCS


Objective: To locate the motor points of the subscapularis muscle in relation to palpable anatomic landmarks and hence suggest a technique for botulinum toxin injection into subscapularis applicable to patients of all ages.

Design: Anatomic dissection of the innervation of 20 subscapular muscles.

Setting: University dissecting room.

Cadavers: Ten formalin-preserved cadavers.

Interventions: Not applicable.

Main Outcome Measure: The location of motor points in relation to anatomic landmarks.

Results: The median number of motor points for the subscapularis was 5 (range, 3–6). All motor point measurements were related to surface points and converted into proportional values along reference lines. Motor points from the 20 dissections showed clustering in a band. A line of best fit was calculated ($y = 1.48x - 0.743$).

Conclusions: We describe an injection technique that would deliver botulinum toxin close to the motor points of the subscapularis, a surrogate marker of the motor endplate zones. By using proportional distances, this technique is applicable to an adult and pediatric population. This should lead to an increased efficacy and decreased side-effect profile in clinical practice, although clinical trials will need to confirm this.

Key Words: Botulinum toxins; Muscle spasticity; Rehabilitation; Shoulder.

© 2007 by the American Congress of Rehabilitation Medicine and the American Academy of Physical Medicine and Rehabilitation

UPPER-LIMB SPASTICITY IS A disabling feature of neurologic conditions such as cerebral palsy and stroke. Spasticity of the subscapularis (along with the other internal rotators) overwhelms the relative weakness of the rest of the rotator cuff resulting in internal rotation and adduction of the arm with associated pain.1

Focal neuronal or neuromuscular blockade has been shown to be effective in the treatment of upper-limb spasticity, relieving pain and improving the range of movement.1-4 Botulinum toxin injection is currently the most popular method employed for achieving this. The toxin blocks transmission at the neuromuscular junction leading to weakness by chemodenervation.5 Many authors believe that maximal efficacy of botulinum toxin is achieved by injecting the toxin as close as possible to its effector site, the motor endplate.6-10

Currently, there is little formal advice on how to inject botulinum toxin into the subscapularis muscle. In general, the simplest method suggested for localizing the site to inject botulinum toxin is palpation of the muscle belly but this is essentially impossible for the subscapularis muscle because of its location. Many clinicians have relied on books such as the Anatomical Guide for the Electromyographer,11 which gives advice on the optimal site of insertion of electrodes for the neurophysiologic study of muscles, to help identify sites for botulinum toxin injection. In this book, there is no mention of the subscapularis muscle. Nerve stimulation and/or formal electromyography to localize the muscle, the motor points, or the motor endplates is also limited in patients with spasticity and dystonias because they are unable to tolerate the procedure without sedation or anesthesia. To identify endplate potentials, the patient must be able to maintain their muscles in a relaxed state.

Motor endplates have been shown to cluster around motor points (the macroscopic entry point of a nerve into a muscle).5,12,13 Therefore, the localization of motor points provides a surrogate marker of the endplate zone and hence the site of injection.

This study investigates the anatomic location of motor points in the subscapularis muscle in relation to surface anatomy landmarks to enable us to suggest a technique for injecting the subscapularis that would be appropriate for use in both children and adults.

METHODS

Twenty subscapular muscles were dissected from 10 cadaveric bodies. The material was obtained under the UK Anatomy Act of 1984. The muscles were skeletally mature and preserved with embalming fluid (methylated spirit, 10% phenol, glycerol, and formalin). Six cadavers were women and 4 were men. The mean age of death ± standard deviation was 74.7 ± 5.8 years (range, 65–84y). There was no evidence of previous injury or surgical procedure around the shoulder joint. The cause of death had not involved any pathologic process that could have affected the innervation pattern of the subscapularis muscle. Cadavers were placed in the prone position, and skin and superficial fascia were removed. The latissimus dorsi and rhomboid major and minor were divided. Serratus anterior was divided at its attachment to the scapula and reflected to reveal the subscapularis muscle. Nerve stimulation and/or formal electromyography to localize the muscle, the motor points, or the motor endplates is also limited in patients with spasticity and dystonias because they are unable to tolerate the procedure without sedation or anesthesia. To identify endplate potentials, the patient must be able to maintain their muscles in a relaxed state.

A metal pin was hammered through the scapula at each motor point and this dorsal position marked. The acromial tip, the medial end of the spine of the scapula, and the inferior angle of the scapula were identified. The dorsal positions of the motor points and surface anatomy positions were recorded onto a transparency overlay and then transferred directly to graph paper.
An x axis was devised such that the vertex of the inferior angle was point (0,0) and the acromial tip always lay on the x axis (fig 1). Coordinates for the medial spine and the motor points were recorded. These coordinates were translated into proportions by dividing the values by the distance along the reference line from the inferior angle of the scapula to the acromial tip. This allowed comparison between specimens. For each dissection, graphical representation of the motor points generated a line of best fit, and the equation of this line was calculated. The average of these 20 lines was then found. The mean position of the medial spine was also calculated. All calculations and graphs were made by using Excel.

RESULTS

The number of motor nerve branches and hence motor points for each subscapularis muscle varied between 3 and 6 with a median value of 5. The distance between the inferior angle and the acromial tip had a mean value of 18.9 cm (range, 16.2–21.1 cm). The scatterplot of all of the motor points from the 20 dissections showed clustering of the motor points in a band (see fig 1). A line of best fit was calculated; this bisected the hypothetical line connecting the inferior angle and the acromial tip (x = 0.50). We believe that this line of best fit represents the optimal path for the injection of botulinum toxin into the subscapularis muscle (fig 2).

DISCUSSION

This study has identified the anatomic motor points of the multipennate subscapularis muscle and related these in terms of proportional distances to 3 fixed bony landmarks. By using this information, we are able to suggest a method for injecting botulinum toxin close to its site of action at the endplate zone within the subscapularis muscle.

Botulinum toxin is taken up by and acts on the terminal nerve endings that form the neuromuscular junction with the motor endplate where it blocks the release of acetylcholine, thus interfering with transmission. Identifying motor endplate zones in large adult human muscles is difficult and involves staining cryosectioned fresh whole muscle. Although in unipennate muscles the motor endplate zones have been shown to cluster in a narrow band around the mid point of the muscle fibers, their location in multipennate muscles such as the subscapularis is less clear. Studies of the gastrocnemius complex in humans have shown that the motor endplate zone lies just distal to the motor points, the proximal limit of the motor endplate zone coinciding with the motor points. Similarly, a study of human facial muscles found the motor endplate zone in the immediate vicinity of the nerves’ entrance point. This confirms current thinking that the terminal arborization of the motor nerve occurs just distal to its point of entry in the muscle. Therefore, it has been suggested that the location of the motor endplates corresponds with the location of the motor points, and hence the motor points are a good surrogate site for injection. It is for these reasons that we elected to identify the motor points.

Botulinum toxin has been shown to be most effective when injected close to the motor endplates. It is able to spread
easily along the muscle length and through muscle fascia\textsuperscript{17} to produce a gradient of chemodenervation, the magnitude and extent of which depends on the dose.\textsuperscript{18} Saturation of a motor endplate zone may occur, and in such cases diffusion of the toxin may be more obvious clinically with effects at distant sites. Hence, most side effects occur when high total doses of botulinum toxin are used.\textsuperscript{19}

The subscapularis is the strongest of the internal rotator muscles of the shoulder and is implicated in spastic pathology as a cause of pain and/or loss of function. The aim of this study was to discover if the motor points of the subscapularis lie in a definable region, which would provide a site for botulinum toxin injection in clinical practice.

Our dissection confirmed the variable nerve supply to the subscapularis as described by Kato.\textsuperscript{20} The superior subscapular nerve usually arises as 2 separate branches (the superior and middle subscapular branches) from the posterior divisions of the upper 2 trunks; these nerves supply the bulk of the muscle. The inferior subscapular nerve gives a branch to the axillary portion or subscapularis and then terminates in and supplies the teres major.\textsuperscript{21}

Despite the variability in the origin and number of nerve branches supplying the subscapularis, our results show that the motor points lie in a constant band across the muscle. The line of best fit shows what we believe to be the optimal path along which to inject botulinum toxin into the subscapularis. Our line of best fit bisects a hypothetical line between the inferior angle of the scapula and its acromial tip; therefore, the needle should be inserted at this point. The angle of injection relative to the line is 56°. This can be achieved by directing the needle to a point approximately half way along the spine of the scapula (see fig 2). To maximize coverage of these motor points, the toxin should be injected from the lateral border of the scapula up until the level of the spine of the scapula.

This technique for injection of the motor points of the subscapularis is best achieved from a posterior axillary fold approach. Chiodo et al\textsuperscript{11} compared 3 methods for needle insertion into the subscapularis (a posterior axillary fold approach, an apical approach, and a medial/vertebral approach) and found the posterior axillary fold approach to be the most needed to define an optimal injection technique for the other internal rotator muscles of the shoulder.

CONCLUSIONS

Identification of the motor points for the subscapularis muscle in relation to surface anatomy landmarks has allowed us to propose a method for the neuromuscular blockade of this muscle that is applicable to both children and adults alike. If injections are delivered more accurately to their site of action, we assume that they will be more clinically effective and that the risk of side effects will be reduced.

Acknowledgments: We thank Ian Johnson, MBChB, Stefan Sadnicki, BSc, and the staff of the University College London dissection room.

References


Supplier

a. Microsoft Corp, One Microsoft Way, Redmond, WA 98052.
Poststroke Shoulder Pain: Its Relationship to Motor Impairment, Activity Limitation, and Quality of Life

John Chae, MD, ME, Don Mascarenhas, MD, David T. Yu, MD, Andrew Kirsteins, MD, Elie P. Elovic, MD, Steven R. Flanagan, MD, Richard L. Harvey, MD, Richard D. Zorowitz, MD, Zi-Ping Fang, PhD


Objective: To assess the relationship between poststroke shoulder pain, upper-limb motor impairment, activity limitation, and pain-related quality of life (QOL).

Design: Cross-sectional, secondary analysis of baseline data from a multisite clinical trial.

Setting: Outpatient rehabilitation clinics of 7 academic medical centers.

Participants: Volunteer sample of 61 chronic stroke survivors with poststroke shoulder pain and glenohumeral subluxation.

Interventions: Not applicable.

Main Outcome Measures: We measured poststroke shoulder pain with the Brief Pain Inventory question 12 (BPI 12), a self-reported 11-point numeric rating scale (NRS) that assesses "worst pain" in the last 7 days. Motor impairment was measured with the Fugl-Meyer Assessment (FMA). Activity limitation was measured with the Arm Motor Ability Test (AMAT) and the FIM instrument. Pain-related QOL was measured with BPI question 23, a self-reported 11-point NRS that assesses pain interference with general activity, mood, walking ability, normal work, interpersonal relationships, sleep, and enjoyment of life.

Results: Stepwise regression analyses indicated that poststroke shoulder pain is associated with the BPI 23, but not with the FMA, FIM, or AMAT scores.

Conclusions: Poststroke shoulder pain is associated with reduced QOL, but not with motor impairment or activity limitation.

Key Words: Pain; Quality of life; Rehabilitation; Shoulder; Stroke.

© 2007 by the American Congress of Rehabilitation Medicine and the American Academy of Physical Medicine and Rehabilitation

Shoulder pain is a common complication of hemiplegia; its reported prevalence ranges between 5% and 84%.1,2 Numerous studies have reported a relationship between poststroke shoulder pain and limited shoulder external rotation range of motion (ROM),3 sensory impairment,4 adhesive capsulitis,5 impingement,6 subluxation,7 spasticity,8 and complex regional pain syndrome (CRPS).9 Its relationship to motor impairment and activity limitation is less clear. Roy et al10 showed that stroke survivors with shoulder pain have more severe motor impairment during recovery. Others, however, have reported no relationship between shoulder pain and motor impairment.6,11 In a study of 108 stroke survivors discharged from a hospital, Wanklyn et al12 reported that patients with shoulder pain had significantly greater activity limitation than patients without pain, based on their Barthel Index of disability scores. Another study,5 however, showed no relationship between shoulder pain and the Barthel Index scores. Finally, other authors2,13 have indicated a relationship between shoulder pain and pain-related quality of life (QOL). This relationship has not been quantitatively demonstrated, however. While the pain experience alone is a sufficient reason for treatment, the importance of treating poststroke shoulder pain is further emphasized if it can be shown that it has a relationship to motor impairment, activity limitation, and QOL.

Our objective in this cross-sectional study was to test the hypothesis that poststroke shoulder pain, motor impairment, activity limitation, and pain-related QOL are statistically related.

METHODS

Participants

We analyzed the baseline data of stroke survivors enrolled in a multicenter randomized clinical trial of percutaneous electric stimulation for the treatment of poststroke shoulder pain.14,15 The clinical trials protocol was approved by the institutional review boards at each participating center. Participants were more than 12 weeks poststroke (hemorrhagic or nonhemorrhagic) and were at least 18 years old. Participants had (1) shoulder pain graded as at least 2 on the 11-point numeric rating scale (NRS) of the Brief Pain Inventory16 question 12 (BPI 12), (2) at least one-half fingerbreadth of inferior glenohumeral separation by palpation with the affected limb in a dependent position.
without manual traction, and (3) cognitive ability to fulfill study requirements (able to recall 3 objects after 30min and use an NRS). Patients were excluded if they had a history of arrhythmia with hemodynamic instability, previous stroke with persistent neurologic deficit, prestroke shoulder pathology, CRPS, any implantable stimulator, or uncontrolled seizures (>1/mo).

**Explanatory Measures**

We defined the primary explanatory or independent variable as being the BPI 12. In studies that predict outcomes, the explanatory or independent variable is the variable that predicts a particular outcome or predicts the dependent variable. The BPI is a multiple-question self-reported metric that assesses both pain intensity (sensory dimension) and the interference (reactive dimension) of pain with daily activities and QOL. The BPI has shown both reliability and validity across cultures and languages. Its developers have suggested that BPI 12, the explanatory pain variable and pain-free external rotation ROM as a secondary stroke specific explanatory variables. Based on results of previous studies, we added pain-free external rotation ROM as a secondary pain–related explanatory measure. Time from stroke onset to study entry, stroke type (hemorrhagic vs nonhemorrhagic), side of hemiplegia, radiographic inferior subluxation, and the Modified Ashworth Scale (MAS) score of the elbow flexors were added as secondary stroke specific explanatory measures. We selected the MAS score of the elbow flexors to avoid the confounding effect of pain associated with shoulder manipulation. The MAS requires that the range be taken to its maximum, limited only by the tone or fixed joint contracture. Many stroke patients, however, have reduced range of shoulder abduction and external rotation because of pain with or without increased tone. Thus, assessment of the shoulder is unlikely to provide an accurate assessment of tone. We selected the elbow on the assumption that tone is generalized.

**Dependent Measures**

Motor impairment and activity limitations were assessed with laboratory-based measures. Motor impairment was assessed with the upper-limb component of the Fugl-Meyer Assessment (FMA). We used 2 measures of activity limitation. Upper-limb specific activity limitation was assessed with the self-care component of the FIM instrument and the functional domain of the Arm Motor Ability Test (AMAT). Pain-related QOL was assessed with BPI question 23 (BPI 23), a self-reported pain questionnaire. Trained therapists administered all measures.

The upper-limb component of the FMA considers evolving synergy patterns as well as isolated strength, coordination, and hypertonia. The FMA’s reliability and validity have been documented. After rigorously evaluating its measurement properties, Gladstone et al concluded, “Based on the available evidence, the Fugl-Meyer Motor scale is recommended highly as a clinical and research tool for evaluating changes in motor impairment following stroke.”

The self-care component of the FIM instrument assesses upper-limb–dependent tasks such as feeding, dressing, bathing, and toileting. The measure permits compensatory strategies and participants may use the unaffected upper limb to perform the tasks. The FIM instrument’s validity, structure, and stability have been demonstrated.

In contrast to the FIM instrument, the AMAT assesses hemiparetic upper-limb–specific functional tasks and does not permit compensatory strategies. Unilateral tasks are performed with the affected upper limb. Bilateral tasks are performed using (or attempting to use) the dominant extremities in the same roles as before the stroke. The correlation between the functional ability and quality of movement domains of the AMAT is very close to 1. Therefore, as recommended by Kopp et al, we used only the functional ability domain. The AMAT has been shown to be reliable, sensitive, valid, and internally consistent.

We assessed pain-related QOL with the BPI 23, which assesses the degree to which pain interferes with a combination of daily activities that includes general activity, mood, walking ability, normal work, interpersonal relationships, sleep, and enjoyment of life. Interference is assessed on an 11-point NRS, where 0 indicates “does not interfere” and 10 indicates “completely interferes.” The summary score is the mean of all 7 domains. Psychometrics was reported earlier in the context of the BPI 12.

**Analysis**

To evaluate the relationship between the explanatory and dependent variables, we generated 4 stepwise linear regression models corresponding to the 4 dependent measures. Table 1 shows the specific explanatory and dependent variables for each model. All models included BPI 12 as the primary explanatory pain variable and pain-free external rotation ROM as a secondary explanatory pain variable. All models also included time from stroke onset to study entry, stroke type, and side of hemiplegia, radiographic inferior subluxation, and MAS score as secondary stroke specific explanatory variables. Because activity limitation may be influenced by motor impairment, we added the FMA as an additional explanatory variable for the second and third models. We added the FIM and AMAT as additional explanatory variables for the fourth model because activity limitation may influence QOL.

<table>
<thead>
<tr>
<th>Model</th>
<th>Primary Explanatory Variable</th>
<th>Secondary Explanatory Variables</th>
<th>Dependent Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>BPI 12</td>
<td>Pain-free external ROM, MAS, stroke onset to study entry, stroke type, side of hemiplegia, subluxation</td>
<td>FMA</td>
</tr>
<tr>
<td>2</td>
<td>BPI 12</td>
<td>Pain-free external ROM, FMA, MAS, stroke onset to study entry, stroke type, side of hemiplegia, subluxation</td>
<td>FIM</td>
</tr>
<tr>
<td>3</td>
<td>BPI 12</td>
<td>Pain-free external ROM, FMA, MAS, stroke onset to study entry, stroke type, side of hemiplegia, subluxation</td>
<td>AMAT</td>
</tr>
<tr>
<td>4</td>
<td>BPI 12</td>
<td>Pain-free external ROM, FMA, MAS, stroke onset to study entry, stroke type, side of hemiplegia, subluxation</td>
<td>BPI 23</td>
</tr>
</tbody>
</table>
RESULTS

Data were available for all 61 stroke survivors enrolled in the clinical trial. The mean age ± standard deviation (SD) of participants was 58.6 ± 12.1 years, 44% were women, 84% sustained nonhemorrhagic strokes, and 62% had left hemiplegia. Other characteristics of the participants are shown in Table 2.

Table 2: Participant Characteristics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean ± SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time from stroke onset to study entry (mo)</td>
<td>58.4 ± 12.1</td>
</tr>
<tr>
<td>BPI 12 score</td>
<td>7.1 ± 2.3</td>
</tr>
<tr>
<td>BPI 23 score</td>
<td>4.3 ± 2.8</td>
</tr>
<tr>
<td>Subluxation (mm)</td>
<td>7.3 ± 0.5</td>
</tr>
<tr>
<td>Pain-free external ROM (deg)</td>
<td>37.3 ± 21.6</td>
</tr>
<tr>
<td>FMA score</td>
<td>18.7 ± 12.6</td>
</tr>
<tr>
<td>MAS score</td>
<td>1.9 ± 1.2</td>
</tr>
<tr>
<td>AMAT function score</td>
<td>0.0 ± 1.0</td>
</tr>
<tr>
<td>FIM self-care score</td>
<td>30.4 ± 7.8</td>
</tr>
</tbody>
</table>

Table 3 shows the results of the stepwise regression analyses. BPI 12 was not associated with the FMA, but pain-free external rotation ROM and the degree of inferior subluxation were directly and inversely related to the FMA, respectively (model 1). The model explained 15% of the variance in the FMA. None of the explanatory variables was associated with FIM self-care (model 2). BPI 12 was not associated with AMAT function (model 3). AMAT function was directly related to FMA, however, and lower AMAT function scores were associated with the non-hemorrhagic stroke subtype. The model explained 72% of the variance in AMAT function. BPI 23 was directly related to FMA, however, (model 2). BPI 12 was not associated with AMAT function (model 4). None of the other explanatory measures was associated with BPI 23.

DISCUSSION

Poststroke shoulder pain was associated with reduced QOL related to pain. Our study, however, failed to demonstrate a statistical relationship between poststroke shoulder pain and motor impairment and activity limitation.

In an earlier World Health Organization definition, health and QOL reflected the constructs of physical, mental, and social well-being and not merely the absence of disease. Today, however, QOL is generally referred to as a multidimensional construct involving the physical, emotional, functional, and social domains, which allows us to view the impact of disability, illness, or pain on a person as a whole.30 Consistent with social domains, which allows us to view the impact of disability, illness, or pain on a person as a whole.30 Consistent with this construct, in this study we used the BPI 23 to assess the degree to which pain interferes with a combination of activities that include general activity, walking ability, mood, vocation, relationships, sleep, and general enjoyment of life. To interfere with specific tasks does not necessarily mean that shoulder pain prevents the completion of the tasks. The pain may simply make the tasks more difficult by requiring greater investment of emotional and volitional effort. Accordingly, BPI 23 has been used as a QOL measure in cancer31 and in low back pain interventional12 studies.

BPI 12, a self-reported measure of “worst pain” experienced in the 7 days before the assessment, was associated with pain-related QOL, but pain-free external rotation ROM was not. This apparent incongruity may be the result of the distinctiveness of each measure. Pain-free external rotation ROM likely overestimates the severity of daily pain. The pain experience is increased when a shortened muscle is stretched or if soft tissue is impinged between the humeral head and the acromion process on movement. Thus, pain-free external rotation ROM introduces artificial constraints that may not be relevant to participants’ routine daily activities. In contrast, BPI 12 assesses worst pain during the past week without specifying a level of activity so that the pain score is most relevant to participants’ real-life activities. Thus, it is not surprising that the stroke survivors’ perception of pain-related QOL was associated with BPI 12, but not pain-free external rotation ROM.

Contrary to expectations, BPI 12 was not associated with the FMA, FIM, or AMAT scores, for which there are several possible explanations. First, this study only included stroke survivors with poststroke shoulder pain. Previous studies10,12 that reported a relationship between poststroke shoulder pain, motor impairment, and activity limitation assessed stroke survivors with and without shoulder pain. Within this broader population of stroke survivors, those with shoulder pain exhibited greater motor impairment and activity limitations. Our subjects all had significant shoulder pain with severe motor impairments and a mean FMA score of 18.7 out of a maximum of 66. The requirement of subluxation, which we correlated with severity of motor impairment, likely presents a selection bias and may have influenced the relationship between poststroke shoulder pain and motor impairment. Similarly, participants who had severely impaired hemiparetic upper-limb specific activity limitations, with a mean AMAT function score of 1.0 out of a maximum of 5. Thus, it is possible that the presence of pain is more important than its degree in predicting motor impairment and activity limitation. Second, with regard to FIM self-care scores, the lack of a relationship to BPI 12 may be because compensatory strategies were allowed. Because the FIM instrument permits the use of the unaffected limb, pain in the poststroke shoulder may be less relevant. Accordingly, participants’ FIM self-care scores were relatively high, with a mean score of 30.4 out of a maximum of 42. Third, because both the FIM instrument and the AMAT assess task comple-

Table 3: Results of Linear Regression Analyses

<table>
<thead>
<tr>
<th>Model</th>
<th>Dependent Variable</th>
<th>Significant Factor(s)</th>
<th>β</th>
<th>F</th>
<th>R²</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 FMA</td>
<td>Overall model</td>
<td>NA</td>
<td>4.9</td>
<td>.15</td>
<td>.01</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Inferior subluxation</td>
<td>-.28</td>
<td>NA</td>
<td>NA</td>
<td>.023</td>
<td></td>
</tr>
<tr>
<td></td>
<td>External rotation ROM</td>
<td>.26</td>
<td>NA</td>
<td>NA</td>
<td>.037</td>
<td></td>
</tr>
<tr>
<td>2 FIM self-care</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 AMAT function</td>
<td>Overall model</td>
<td>NA</td>
<td>74.6</td>
<td>.72</td>
<td>.&lt;.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>FMA</td>
<td>.83</td>
<td>NA</td>
<td>NA</td>
<td>.&lt;.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Stroke type</td>
<td>-.20</td>
<td>NA</td>
<td>NA</td>
<td>.005</td>
<td></td>
</tr>
<tr>
<td>4 BPI 23</td>
<td>BPI 12</td>
<td>.46</td>
<td>15.4</td>
<td>.20</td>
<td>.&lt;.001</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviation: NA, not applicable.
*No significant factor.
tion, it is possible that participants experienced significant pain but were still able to complete at least some of the prescribed tasks. Finally, the FMA, FIM instrument, and AMAT assess performance in the laboratory whereas BPI 12 and 23 reflect real-life experiences and activities, respectively. Participant performance in the laboratory and in real-life activities may not be congruent. Specifically, the incentive to complete a task is higher in the laboratory than at home, where some factors may actually provide disincentives. A factor such as pain, which may inhibit but not preclude certain tasks, has a greater effect in lower incentive environments.

CONCLUSIONS

Poststroke shoulder pain is associated with pain-related QOL. This provides a further incentive to develop effective rehabilitation prevention and treatment strategies for poststroke shoulder pain. Our data did not identify a relationship between poststroke shoulder pain, motor impairment, and activity limitation. This, however, may have been an artifact of a study design that did not include stroke survivors free of poststroke shoulder pain, rather than a true lack of a relationship.

References

Measurement Properties of the National Institutes of Health Stroke Scale for People With Right- and Left-Hemisphere Lesions: Further Analysis of the Clomethiazole for Acute Stroke Study–Ischemic (Class-I) Trial

Scott R. Millis, PhD, Don Straube, MS, PT, Cherdask Iramaneerat, MD, Everett V. Smith Jr, PhD, Patrick Lyden, MD


Objective: To assess the psychometric properties of the National Institutes of Health Stroke Scale (NIHSS) in people with either left or right acute hemisphere stroke for the purpose of improving the scale’s sensitivity in detecting neurologic impairment.

Design: Secondary analysis of data from the Clomethiazole for Acute Stroke Study–Ischemic using the Rasch partial credit model. We evaluated the data’s measurement properties using item-total correlations, Rasch item fit statistics, principle component analysis of standardized person and item residuals, differential item functioning, separation reliability, and the separation ratio.

Setting: Original data were collected in academic and community hospitals as part of a clinical trial.

Participants: People with acute ischemic stroke who were seen within 12 hours of onset: 380 people with left-hemisphere stroke and 347 with right-hemisphere stroke.

Interventions: Not applicable.

Main Outcome Measure: The NIHSS.

Results: Items of the NIHSS function differently in the right- and left-hemisphere lesion groups. We constructed for each group separate linear scales consisting of a subset of items of the NIHSS to improve its measurement properties.

Conclusions: Our findings provide initial support for the use of individual, targeted scales for measurement of impairment after ischemic stroke. Low person separation reliability may be a consequence of the sample, which included only people with large ischemic cortical strokes.

Key Words: Cerebrovascular accident; Outcome assessment (health care); Psychometrics; Rehabilitation.

© 2007 by the American Congress of Rehabilitation Medicine and the American Academy of Physical Medicine and Rehabilitation

The National Institutes of Health Stroke Scale (NIHSS) is among the most widely used instruments for the evaluation of neurologic impairment after stroke and has recently been recommended for use in assessing acute stroke by the Stroke Council of the American Heart Association. The NIHSS has 15 items that evaluate several aspects of neurologic function, including level of consciousness, eye movement, visual field deficit, language and speech, and motor and sensory involvement. The 15 items were selected on the basis of the empirical literature and expert opinion. The scale has acceptable interrater reliability among neurologists and other clinicians, and is also substantially correlated with the Barthel Index of activity of daily living and the Orpington Prognostic Scale, as well as lesion volume as measured by computed tomography and magnetic resonance imaging. It has been found to be predictive of hospital length of stay, hospital discharge destination, and survival.

Some NIHSS items, however, have been shown to be unreliable, redundant, or not contributory to the internal structure in a meaningful way (e.g., level of consciousness, ataxia, facial weakness, and dysarthria). Studies have indicated that the instrument could be successfully shortened by eliminating these items while remaining a valid and reliable measure of stroke severity. In addition, factor analyses have suggested that the NIHSS has 2 underlying constructs that represent left and right cortical and motor function. Although both factors represent cerebral function, people with right- or left-hemispheric strokes could constitute 2 distinct patient populations. Separate scores for different populations addressed by each factor may be required to properly evaluate a heterogeneous patient group with a single instrument. For example, Woo et al found that the NIHSS may be more sensitive to left-hemisphere lesions than to right-hemisphere lesions because it has more language items than items that assess functions that are predominately mediated by the right hemisphere. Hence, the NIHSS may tend to overestimate the severity of stroke in the left hemisphere or, conversely, underestimate it in right-hemisphere lesions. The same score may have different implications, depending on lesion location.

Ideally, instruments should target just 1 attribute or construct, referred to as unidimensionality. In addition, interval scale measures are needed for meaningful comparisons of illness or injury severity using parametric statistics. The usual practice of summing the ratings across all NIHSS items and treating the total score as an interval scale metric is problematic, not only because the items are ordinal but because the scale may not be sufficiently unidimensional. If that should
be the case, a preferred approach would be to derive 2 separate scales with which to assess the distinct effects on function that are associated with lesion location and convert these ordinal measures into interval measures. This could also make the NIHSS a more targeted and efficient measure, which was our purpose in this investigation. Such complexity, however, may not significantly improve the real-world use of the NIHSS. Therefore, we sought to test whether the scale detects sufficient differences in right- versus left-hemisphere strokes to justify clinical reporting of each hemispheric score separately.

METHODS

Participants

Participants in this study were patients enrolled in the Clomethiazole for Acute Stroke Study–Ischemic (CLASS-I) clinical trial. The CLASS-I trial methods and results have been previously published. The current investigation was a secondary analysis of the deidentified database containing the NIHSS data from the CLASS-I clinical trial. This study was reviewed and was qualified as an exempt protocol by the Wayne State University Human Investigation Committee. The institutional review board at the University of Illinois at Chicago also approved the protocol. Participants had a major ischemic stroke and a combination of limb weakness, higher cortical dysfunction, and visual field deficits within 12 hours of symptom onset. Mean patient age ± standard deviation (SD) was 72 ± 12 years; 53% were women. Medical histories included habitual smoking by 18% of the subjects, diabetes mellitus in 24%, current hypertension in 75%, and prior stroke in 20%. Mean and median NIHSS raw scores were 17. The original sample included 1200 patients; however, for the purposes of this study, we included only 380 patients with left-hemisphere stroke and 347 patients with right-hemisphere stroke. We excluded 40 patients with brainstem localization, 426 with no localization specified, and 7 who had no recorded NIHSS scores.

Instrument and Procedure

Participants were assessed with the NIHSS within 12 hours of symptom onset. CLASS-I investigators underwent NIHSS certification with training and testing videotapes. Recertification was conducted every year.

Data Analysis

Rasch analysis, based on the mathematical model formulated by Rasch, can be used to construct and evaluate measurement scales. We used Rasch analysis in this investigation to estimate the degree of each subject’s neurologic impairment and the severity of that impairment as associated with each NIHSS item on a common logit scale. The Rasch model is a probabilistic model that converts the ordinal scores derived by summing item scores into interval measures when the data fit the model’s expectations. Rasch analysis can be used to assess several psychometric properties, including unidimensionality (the degree to which the items assess 1 construct); item difficulty (ordering items from least to most difficult to perform); and targeting (the extent to which items are of appropriate difficulty for a given group of patients). Items of the NIHSS are scored on a 3- to 5-point ordinal scale, with higher scores indicating more severe neurologic impairment. We used the Rasch partial credit model in this study because the scale’s items allow for a degree of correct or “impaired” response and the level of impairment varies from 1 item to the next. In addition, we used a process known as “pivot anchoring” to create comparable rating scales from the differently worded and defined categories of the NIHSS items. More precisely, pivot anchoring is a way to align rating scale categories so that the point chosen to define impairment is comparable across the items. In the case of the NIHSS, this would be from the response categories of 0 to 1 (ie, performs the task correctly vs impaired performance) for each of the 15 items.

We checked the internal consistency of the stroke scores of patients over different items and of items over different patients using separation reliability. The separation reliability of patients indicates how consistent their stroke scores are across items. Conversely, the separation reliability of items indicates how consistent the item scores are across patients. The separation reliability is an estimate of the ratio of true score variance to observed score variance. Its value can range from 0 to 1. Because of this limited range, the separation reliability may suffer from ceiling effect. To avoid this potential effect, we also report the separation ratio associated with each separation reliability. The separation ratio represents the ratio of the true SD of measures over the average standard error of measures and can range from 0 to infinity.

We performed the initial analysis with the entire sample so as to examine the scale structure based on the total sample. If there was a departure from unidimensionality, the scale structure of the right- and left-hemisphere lesion groups would be examined separately and linear scales would be generated for each group. We used the Rasch analysis program Winsteps in all analyses.

We used 4 methods to examine the dimensionality of the NIHSS data. First, because we hypothesized that there would be 2 different populations with corresponding different internal structures, we performed Rasch principal components analyses (PCA) of standardized person and item residuals. The ideal structure for a Rasch analysis is for all the information in the data to be explained by a single latent trait, which in the case of the NIHSS, is neurologic impairment associated with stroke. The unexplained portion of the data, or residual variance, is assumed to be random noise. The aim of PCA of Rasch residuals is to extract the common factor that explains the most residual variance of the items and person measures under the hypothesis that there may be additional factors.

Second, we analyzed item fit statistics to assess the contribution of each item of the NIHSS to severity of stroke impairment. The Winsteps program provides 2 types of fit statistics for items and persons: both are based on the chi-square statistic. Infit statistics are sensitive to unexpected responses (violations of the model assumptions) near a subject’s impairment level. Outfit statistics are sensitive to unexpected responses far from a subject’s symptom level. The fit statistics, when reported as mean squares, have an expected value of 1 and range from 0 to infinity. Values less than 1 may indicate a lack of expected stochasticity in the data. Values greater than 1 indicate departures from unidimensionality. For rating scale data, a reasonable range for mean-square fit statistics is 0.6 to 1.4 and was the criterion we used in this investigation. Although the Rasch model fit statistics can indicate the extent of misfit of an item, they do not explain why the item misfits. One possibility is that the item is not consistent with the underlying construct of interest.

Third, given the hypothesis that 2 different populations exist (right- and left-hemisphere lesion groups), we also investigated the functioning of the items across these groups, using the differential item functioning (DIF) as implemented in Winsteps. The DIF evaluation essentially entails determining whether the level of neurologic severity represented by each
item is the same, within measurement error, across the 2 population groups. Finally, for the combined sample, we examined item-total correlations for each item as another check on whether any items tap into a construct that is different from other items.

RESULTS

The initial analysis of all 15 items with the entire sample yielded a person separation ratio of 0.69 with a person reliability of 0.32, and an item separation ratio of 30.31 with an item reliability of 1.00. Item measures ranged from $-3.31$ (visual fields) to 4.04 (ataxia) logits. The analysis revealed that several items did not fit with model expectations. Ataxia, left arm, and left leg items had high infit and outfit mean-square statistics, which suggested that they might measure a construct that is different from other items or represent 2 different levels of severity for multiple subpopulations within the overall sample.

PCAs of standardized item and person residuals supported the existence of 2 distinct subject populations and measurement scales. In the PCA of standardized item residuals, the first factor extracted 35.33% of item residual variance. The items separated into 2 dimensions. The first group, which had positive loadings, had 6 items: language, right arm, questions, right leg, commands, and dysarthria. These items tend to represent the impairment in the functions mediated by the dominant cerebral hemisphere, which is the left hemisphere in 90% of all people.26 The second group, which had negative loadings, had 9 items: left leg, left arm, extinction/inattention, sensory, gaze, facial palsy, visual fields, level of consciousness, and ataxia. These items clustered into 2 subgroups: 1 with large negative loadings and 1 with low to moderate negative loadings. Left leg and arm items, which had large negative loadings, assess motor functions associated with the contralateral right hemisphere. In addition, neglect (extinction/inattention) may be more commonly associated with right-hemisphere lesions.31 The remaining items with low-to-moderate negative loadings (eg, visual fields, sensory, gaze, facial palsy) may have appeared with this second group as a consequence of how some items are scored when patients are aphasic.31 That is, clear cut deficits must be demonstrated for the sensory and visual fields and facial palsy items if they are to be scored as impaired. Hence, it is not surprising that these language-dependent items loaded on the right-hemisphere dimension because many of the left-hemisphere lesion patients with aphasia could not perform these items correctly and consequently were scored as unimpaired, per NIHSS scoring guidelines.

Following the extraction of the primary dimension of neurologic impairment, the PCA of person residuals identified 2 populations. Among 378 subjects with positive component loadings, only 7 (1.85%) had right-hemisphere stroke. Conversely, among 349 subjects with negative component loadings, only 9 (2.58%) had left-hemisphere stroke. The first factor extracted 41.47% of person residual variance. This pattern of findings provided further evidence that after extracting the best common variable, the participants were further differentiated by lesion location.

To determine whether the NIHSS items had significantly different meanings for the right-hemisphere and left-hemisphere stroke groups, DIF was conducted with Winsteps.26 The results of the DIF analysis indicated that 14 of the 15 items represented statistically different ($P < .001$) severities of neurologic impairment between people with right-hemisphere stroke and subjects with left-hemisphere stroke. Ataxia was the only item not displaying DIF. Finally, the ataxia and level of consciousness items correlated poorly with the overall scale, a

In summary, findings from the PCAs of the standardized item residuals suggest that the items in the NIHSS measure 2 separate dimensions of neurologic impairment in stroke. The first dimension is associated with items indicating lesions of the left hemisphere and the second is associated with items indicating right- or left-hemisphere lesions. The PCAs of the person residuals and the DIF analysis suggest that the overall population may be comprised of 2 distinctive subpopulations based on lesion location. These findings supported additional analyses of the scale properties of the NIHSS by separating the sample into lesion groups.

Left-Hemisphere Stroke Group

When we limited our analysis to the 380 participants with left-hemisphere stroke, using all 15 items, the NIHSS yielded a person separation ratio of 1.64 with a person reliability of 0.73 and an item separation ratio of 23.43 with an item reliability of 1.00. This subgroup analysis produced a more reliable measure that separated participants according to their level of impairment better than a measure that analyzed all participants together. The person reliability, which is analogous to the Cronbach $\alpha$, increased from 0.32 to 0.73. Item measures ranged from 3.02 logits (visual field) to 5.01 logits (ataxia). The 3 misfitting items were left arm, ataxia, and left leg (table 1). We increased the precision (ie, person separation and reliability) of the measure and fit of the data by eliminating these 3 items from the scale. An analysis of subjects with left-hemisphere stroke after the 3 misfitting items were eliminated yielded a person separation ratio of 1.79 with a person reliability of 0.76 and an item separation ratio of 15.28, with an item reliability of 1.00. Item measures ranged from $-2.11$ logits (visual fields) to 2.74 (level of consciousness) (see table 1). All item mean-square fit statistics were within the 0.6 to 1.4 criteria.

The variable map of subjects and items (fig 1) shows that these items reasonably target the subjects’ neurologic deficits. Examination of the item hierarchy reveals that for both left-hemisphere and right-hemisphere group, the most common symptoms were visual field defect, facial palsy, right-sided motor impairment, and language impairment. In interpreting figure 1, it should be emphasized that items higher up in the variable map represent a higher degree of neurologic impairment. For example, unresponsiveness on the level of consciousness item reflects greater impairment than facial palsy. The 12 items in the scale target much of the level of neurologic impairment exhibited by the participants with minimal redundancy of items. Additional items targeted at the higher end of the severity continuum may be desirable, however.

Right-Hemisphere Stroke Group

When the analysis was limited to the 347 participants with right-hemisphere stroke, again using all 15 items, the NIHSS yielded a person separation ratio of 1.27 with a person reliability of 0.62 and an item separation ratio of 25.74, with an item reliability of 1.00. Analysis of this subgroup found that separating participants according to their levels of neurologic impairment yielded a more reliable measure than when all subjects were analyzed together. Item measures range from $-3.91$ logits (visual fields) to 4.55 logits (ataxia). Not surprisingly, the most misfitting items were ataxia, right arm, right leg, and language (table 2).

High outfit mean-square values of ataxia, right arm, right leg, and language items suggest that these items had excessive error variance. We eliminated the 3 most misfitting items (ie, ataxia, right arm, right leg) to improve the reliability and fit of the data and re-ran the analysis. This analysis yielded a person
separation ratio of 1.38 with a person reliability of .66 and an item separation ratio of 23.02, with an item reliability of 1.00. Item measures range from −3.01 logits (visual fields) to 5.99 logits (commands) (see Table 2). The variable map of subjects and items (Fig 2) shows that these items target well the neurologic impairment of this subgroup. As with the left-hemisphere group, the most common deficits involved contralateral motor functions and facial palsy. Language impairment, however, is relatively uncommon in right-hemisphere stroke and, if present, indicates significant neurologic impairment: the language item is highest in the hierarchy.

**DISCUSSION**

Our Rasch analysis found evidence that the NIHSS measures 2 distinct constructs, which confirms the findings of previous factor analytic studies.1,14 These 2 constructs reflect neurobehavioral functions differentially mediated by the left and right hemispheres, respectively. Moreover, our initial findings support the use of individual, targeted scales for measurement of impairment after stroke.1,16 When the data from the total sample were analyzed, the PCA of standardized item and person residuals demonstrated 2 groups of items unique to people with left- and right-hemisphere stroke, as well as 2 groups of participants separated by the side of their lesions. In addition, the DIF analysis demonstrated that all items except ataxia represent different degrees of neurologic impairment for patients with right-hemisphere stroke and those with left-hemisphere stroke. Finally, when the analysis was performed on each group separately, scales were generated that demonstrated that the items of the NIHSS function differently across the 2 groups.

Given the contralateral organization of motor functions, it is not surprising that the right arm and leg tasks of the NIHSS demonstrated good fit for the scale for the left-hemisphere stroke group, but not for right-hemisphere group. Similarly, the items related to left arm and leg tasks demonstrated good fit for the right-hemisphere group, but not for the left-hemisphere group. In addition, the language and command items fit the scale for the left-hemisphere group, which provides evidence of their contribution to the scale’s validity.

By contrast, these language and command items functioned differently for a subset of participants in the right-hemisphere group. That is, some subjects responded to these items in a manner inconsistent with the modeled expectation for their level of neurologic impairment. The “language” item assesses both fluency and comprehension and is multifaceted: patients are asked to describe what is occurring in a picture, to name items, and to read sentences. In addition, comprehension is rated not only on how patients respond to the language item, but to all of the commands in the preceding general neurologic examination. Thus, some participants may have responded to this item in a unique manner that contributed to the high outfit mean-square statistic. The command item may be problematic for a different reason, that is, participants can respond to the item in various ways: (1) opening and closing the eyes or the nonparetic hand on verbal command, (2) substituting another 1-step command if the hands cannot be used, or (3) pantomiming the task as demonstrated by the examiner. These different methods potentially confound the fit of the command item because they may not be equivalent. It is not possible to know which version of the item the examiner used when administering the NIHSS, and thus it is difficult to fully explain why the items demonstrated high outfit mean-square values.

The original version of the NIHSS included only the involved limbs in the motor items (item 5 for arm, item 6 for leg).2 This approach proved effective during small pilot clinical trials and obviously avoids the confounding factor that occurred when we included right- and left-hemisphere patients in 1 scale. Later investigators3 added both limbs (expanding item 5 to include 5a and 5b, and item 6 to include 6a and 6b, to indicate the left and right sides, respectively) when the scale was used in larger trials that included a greater variety of strokes. The modification improved the specificity of the data collected from the study sites,3 but introduced the problem of heterogeneity into the scale. As data collection in health care moves from paper and pencil formats to computerized formats (computer adaptive testing), use of screening items based on patient diagnosis should allow for selection of more targeted items while simultaneously sim-

<table>
<thead>
<tr>
<th>Items</th>
<th>Measures (logits)</th>
<th>Infit Mnsq</th>
<th>Outfit Mnsq</th>
<th>Measures (logits)</th>
<th>Infit Mnsq</th>
<th>Outfit Mnsq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ataxia</td>
<td>5.01</td>
<td>1.19</td>
<td>9.90</td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Left arm</td>
<td>3.90</td>
<td>1.15</td>
<td>9.90</td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Left leg</td>
<td>3.06</td>
<td>1.15</td>
<td>4.71</td>
<td>2.74</td>
<td>1.19</td>
<td>1.32</td>
</tr>
<tr>
<td>Consciousness</td>
<td>1.53</td>
<td>1.06</td>
<td>1.10</td>
<td>1.61</td>
<td>1.06</td>
<td>1.03</td>
</tr>
<tr>
<td>Commands</td>
<td>0.54</td>
<td>1.02</td>
<td>1.01</td>
<td>0.83</td>
<td>1.12</td>
<td>1.11</td>
</tr>
<tr>
<td>Extinction</td>
<td>−0.22</td>
<td>1.03</td>
<td>1.01</td>
<td>0.72</td>
<td>0.93</td>
<td>1.14</td>
</tr>
<tr>
<td>Questions</td>
<td>−0.23</td>
<td>0.92</td>
<td>0.94</td>
<td>0.71</td>
<td>0.91</td>
<td>0.91</td>
</tr>
<tr>
<td>Gaze</td>
<td>−0.37</td>
<td>0.85</td>
<td>0.84</td>
<td>0.45</td>
<td>1.08</td>
<td>1.24</td>
</tr>
<tr>
<td>Dysarthria</td>
<td>−0.85</td>
<td>0.99</td>
<td>1.08</td>
<td>−0.21</td>
<td>1.10</td>
<td>1.10</td>
</tr>
<tr>
<td>Sensory</td>
<td>−1.25</td>
<td>1.05</td>
<td>1.05</td>
<td>−0.48</td>
<td>0.81</td>
<td>0.76</td>
</tr>
<tr>
<td>Language</td>
<td>−1.38</td>
<td>0.80</td>
<td>0.75</td>
<td>−0.48</td>
<td>0.81</td>
<td>0.76</td>
</tr>
<tr>
<td>Right leg</td>
<td>−1.93</td>
<td>0.85</td>
<td>0.81</td>
<td>−1.61</td>
<td>0.87</td>
<td>0.80</td>
</tr>
<tr>
<td>Right arm</td>
<td>−2.50</td>
<td>0.88</td>
<td>0.81</td>
<td>−1.64</td>
<td>0.97</td>
<td>0.98</td>
</tr>
<tr>
<td>Facial palsy</td>
<td>−2.57</td>
<td>0.90</td>
<td>0.91</td>
<td>−2.11</td>
<td>1.02</td>
<td>0.96</td>
</tr>
<tr>
<td>Visual field</td>
<td>−3.02</td>
<td>1.00</td>
<td>0.95</td>
<td>0.00</td>
<td>1.00</td>
<td>1.02</td>
</tr>
<tr>
<td>Mean</td>
<td>0.00</td>
<td>0.99</td>
<td>2.39</td>
<td>1.39</td>
<td>0.11</td>
<td>0.16</td>
</tr>
<tr>
<td>SD</td>
<td>2.34</td>
<td>0.12</td>
<td>3.09</td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

NOTE. Fit statistics not meeting fit criteria are in boldface. Abbreviation: Mnsq, mean square.

*Not included in the second model.
plifying instrument use and facilitating data collection processes.

Although many of the items fit both subject populations and scales, the items functioned differently across the 2 groups. This fact was supported by the DIF analysis and can be visualized by comparing the item hierarchy in figures 1 and 2. The rank order of the items was different between lesion groups and thus, comparison of raw scores between groups may not convey the same meaning. For example, visual field, motor, and sensory impairments are commonly endorsed items in both groups. The language item is higher in the hierarchy in the right-hemisphere group, however, which indicates greater neurologic impairment when endorsed in this group. By contrast, language impairment is commonly found in the left lesion group, which is consistent with the neurobehavioral organization of the brain. In addition, Weintraub and Mesulam, among others, have argued that the right hemisphere is dominant for attentional functions in a manner similar to the left hemisphere’s dominance for language. Indeed, the “extinction/inattention” item was more frequently impaired in the right-hemisphere group than in the left-hemisphere group. Clearly, interpretations of the meaning of total raw scores differ depending on the population to which one belongs.

Consistent with previous studies, we found that the ataxia item was poor fitting when the entire sample was analyzed, as well as when the sample was split into lesion groups and reanalyzed. Consideration might be given to eliminating that item. Still, some investigators have found the ataxia item useful in the assessment of brainstem strokes and have suggested retaining it for routine clinical assessment.

Our findings in this study have several implications for clinical practice involving acute stroke patients. First, we do not advocate dropping the standard NIHSS summary score. Rather, we suggest that the standard score be supplemented with the Rasch-transformed subtest score corresponding to side of lesion until there is more evidence with which to substantiate the clinical utility of the current findings. Second, the results of this study may also help identify people with mild impairments associated with acute right cortical stroke. Foerch et al argued that these subjects are often underdiagnosed compared with those with acute left cortical lesions, particularly when people present with mild neurologic impairment. This is not surprising considering that previous studies have suggested that the items of the NIHSS emphasize deficits associated with left-hemispheric lesions. We found that the NIHSS items that target the

<table>
<thead>
<tr>
<th>PERSONS</th>
<th>ITEMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;more impaired&gt;</td>
<td>&lt;tasks on which only most impaired persons have difficulty&gt;</td>
</tr>
<tr>
<td>5</td>
<td>+</td>
</tr>
<tr>
<td>4</td>
<td>### +</td>
</tr>
<tr>
<td>3</td>
<td>### +</td>
</tr>
<tr>
<td>2</td>
<td>.#####</td>
</tr>
<tr>
<td>1</td>
<td>.#####</td>
</tr>
<tr>
<td>0</td>
<td>.##</td>
</tr>
<tr>
<td>-1</td>
<td>.##</td>
</tr>
<tr>
<td>-2</td>
<td>.##</td>
</tr>
<tr>
<td>-3</td>
<td>&lt;less impaired&gt;</td>
</tr>
</tbody>
</table>

Fig 1. Variable map of level of neurologic deficit of subjects with left-hemisphere stroke and item difficulty level (in logits). Abbreviations: M, mean; S, 1 standard deviation; T, 2 standard deviations. Legend: each “#” represents 3 subjects; each “.” represents 1 subject.
least impaired people with acute right hemispheric stroke were visual fields, left arm, left leg, facial palsy, and extinction/inattention. Third, given the amount of evidence that supports the use of separate measurement scales based on lesion location, application of the results from the separate calibrations might help clinicians characterize more accurately levels of impairment and help guide treatment and disposition.

Study Limitations

Our study has some limitations. The CLASS-I trial targeted people with large cortical strokes: subjects with subcortical stroke

### Table 2: Item Statistics for Participants With Right-Hemisphere Stroke

<table>
<thead>
<tr>
<th>Items</th>
<th>Measures (logits)</th>
<th>Infit Mnsq</th>
<th>Outfit Mnsq</th>
<th>Measures (logits)</th>
<th>Infit Mnsq</th>
<th>Outfit Mnsq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ataxia</td>
<td>4.55</td>
<td>1.09</td>
<td>6.16</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Right arm</td>
<td>3.95</td>
<td>1.14</td>
<td>3.92</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Right leg</td>
<td>3.28</td>
<td>1.02</td>
<td>1.49</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Commands</td>
<td>2.91</td>
<td>0.94</td>
<td>1.04</td>
<td>3.99</td>
<td>0.96</td>
<td>1.48</td>
</tr>
<tr>
<td>Language</td>
<td>2.15</td>
<td>1.01</td>
<td>1.47</td>
<td>3.21</td>
<td>1.06</td>
<td>2.28</td>
</tr>
<tr>
<td>Questions</td>
<td>1.62</td>
<td>0.94</td>
<td>0.96</td>
<td>2.67</td>
<td>1.01</td>
<td>1.30</td>
</tr>
<tr>
<td>Consciousness</td>
<td>0.96</td>
<td>1.01</td>
<td>0.99</td>
<td>2.02</td>
<td>1.02</td>
<td>1.01</td>
</tr>
<tr>
<td>Gaze</td>
<td>-0.68</td>
<td>0.88</td>
<td>0.87</td>
<td>0.31</td>
<td>0.91</td>
<td>0.91</td>
</tr>
<tr>
<td>Dysarthria</td>
<td>-1.31</td>
<td>0.96</td>
<td>0.95</td>
<td>-0.31</td>
<td>0.98</td>
<td>0.98</td>
</tr>
<tr>
<td>Sensory</td>
<td>-1.88</td>
<td>1.14</td>
<td>1.18</td>
<td>-0.91</td>
<td>1.16</td>
<td>1.19</td>
</tr>
<tr>
<td>Extinction</td>
<td>-2.28</td>
<td>0.98</td>
<td>1.08</td>
<td>-1.33</td>
<td>0.99</td>
<td>1.13</td>
</tr>
<tr>
<td>Facial palsy</td>
<td>-2.71</td>
<td>0.99</td>
<td>1.02</td>
<td>-1.79</td>
<td>1.00</td>
<td>1.04</td>
</tr>
<tr>
<td>Left leg</td>
<td>-3.12</td>
<td>0.85</td>
<td>0.87</td>
<td>-2.21</td>
<td>0.82</td>
<td>0.89</td>
</tr>
<tr>
<td>Left arm</td>
<td>-3.53</td>
<td>0.85</td>
<td>0.81</td>
<td>-2.64</td>
<td>0.78</td>
<td>0.71</td>
</tr>
<tr>
<td>Visual field</td>
<td>-3.91</td>
<td>0.99</td>
<td>0.97</td>
<td>-3.01</td>
<td>1.01</td>
<td>1.01</td>
</tr>
<tr>
<td>Mean</td>
<td>0.00</td>
<td>0.99</td>
<td>1.59</td>
<td>0.00</td>
<td>0.98</td>
<td>1.16</td>
</tr>
<tr>
<td>SD</td>
<td>2.82</td>
<td>0.09</td>
<td>1.43</td>
<td>2.31</td>
<td>0.10</td>
<td>0.39</td>
</tr>
</tbody>
</table>

**NOTE.** Fit statistics not meeting fit criteria are in boldface.

*Not included in the second model.

Fig 2. Variable map of level of neurologic deficit of subjects with right-hemisphere stroke and item difficulty level (in logits). Legend: each "#" represents 3 subjects; each "*" represents 1 subject.
or small cortical strokes were not represented. In addition, because only ischemic strokes were included, our findings may not apply to people with hemorrhagic strokes. Consequently, these sample characteristics may have also contributed to the relatively low person separation reliability estimates. Although person separation reliability increased substantially from .32 for the full sample to .76 and .66 for the left- and right-hemisphere groups, respectively, these levels may not be sufficient for some types of research (eg, studies of individual differences). Future studies of the NIHSS should include a more diverse sample of stroke patients so that this hypothesis can be evaluated. Moreover, pending acquisition of more evidence substantiating the clinical utility of the current findings, we are not advocating that clinicians stop using the original NIHSS Scale and interpretation guidelines. Future studies of the NIHSS should also include continued assessment of item psychometric properties and predictive validity studies.

CONCLUSIONS

Our findings in this study support the use of group-specific, unidimensional scales for the assessment of impairment after acute left and right cortical stroke. The subscales that we derived from the original pool of NIHSS items are consistent with current conceptualizations of neurobehavioral functions mediated by each hemisphere. Items that did not contribute to unidimensionality were eliminated, resulting in scales that are more targeted and may be more efficient in clinical practice. Nonetheless, there is still low person separation reliability, which is a consequence of a sample that consisted only of people with large ischemic cortical strokes.

References

18. Wright BD, Linacre JM. Observations are always ordinal; measurements, however, must be interval. Arch Phys Med Rehabil 1989;70:857-60.

Supplier
a. Winsteps, PO Box 811322, Chicago IL 60681.
Clinimetric Properties of the Duruoz Hand Index in Patients With Stroke

Nebahat Sezer, MD, Gunes Yavuzer, MD, Koncay Sivrioglu, MD, Pınar Basaran, MD, B. Fusun Koseoglu, MD


Objective: To investigate the reliability, validity, and responsiveness of the Duruoz Hand Index (DHI) in assessing activity limitation related to hand function in patients with stroke.

Design: Prospective validation study. A consecutive sample of stroke patients was evaluated on 3 occasions: 2 baseline measurements with a 24-hour interval in between, and again 1 month later immediately after a 4-week inpatient rehabilitation program.

Setting: Three different inpatient rehabilitation centers.

Participants: A consecutive sample of 56 patients with stroke (33 men, 23 women) with a mean age 62 years and a mean time since stroke 84 days.

Interventions: Not applicable.

Main Outcome Measures: Brunstrom stages, Modified Ashworth Scale, sensory status, FIM instrument, and DHI. Test-retest reliability was tested using the intraclass correlation coefficient (ICC) and internal consistency was tested using the Cronbach α coefficient. Indexes of measurement error were calculated by standard error of measurement and minimal detectable change (MDC). Construct validity was assessed by association with the FIM instrument (Spearman ρ correlation coefficient). Responsiveness was assessed by calculation of the effect size and paired t test.

Results: The test-retest reliability and internal consistency of the DHI were excellent, with an ICC of .99 (95% confidence interval, .93-.99) and α of .97. The MDC was 1.4 DHI points. The correlation between the DHI and the FIM self-care items was high (r = −.73). The DHI significantly discriminated the patients with dominant side paresis versus nondominant side paresis (P < .01). The DHI score improved significantly after a 4-week inpatient rehabilitation program (P < .05).

Conclusions: The DHI is a time and labor efficient, practical instrument that can be used to assess the hand-related activity level for clinical and research purposes in patients with stroke.

Key Words: Cerebrovascular accident; Hand; Outcome assessment (health care); Rehabilitation.

© 2007 by the American Congress of Rehabilitation Medicine and the American Academy of Physical Medicine and Rehabilitation


Objective: To investigate the reliability, validity, and responsiveness of the Duruoz Hand Index (DHI) in assessing activity limitation related to hand function in patients with stroke.

Design: Prospective validation study. A consecutive sample of stroke patients was evaluated on 3 occasions: 2 baseline measurements with a 24-hour interval in between, and again 1 month later immediately after a 4-week inpatient rehabilitation program.

Setting: Three different inpatient rehabilitation centers.

Participants: A consecutive sample of 56 patients with stroke (33 men, 23 women) with a mean age 62 years and a mean time since stroke 84 days.

Interventions: Not applicable.

Main Outcome Measures: Brunstrom stages, Modified Ashworth Scale, sensory status, FIM instrument, and DHI. Test-retest reliability was tested using the intraclass correlation coefficient (ICC) and internal consistency was tested using the Cronbach α coefficient. Indexes of measurement error were calculated by standard error of measurement and minimal detectable change (MDC). Construct validity was assessed by association with the FIM instrument (Spearman ρ correlation coefficient). Responsiveness was assessed by calculation of the effect size and paired t test.

Results: The test-retest reliability and internal consistency of the DHI were excellent, with an ICC of .99 (95% confidence interval, .93-.99) and α of .97. The MDC was 1.4 DHI points. The correlation between the DHI and the FIM self-care items was high (r = −.73). The DHI significantly discriminated the patients with dominant side paresis versus nondominant side paresis (P < .01). The DHI score improved significantly after a 4-week inpatient rehabilitation program (P < .05).

Conclusions: The DHI is a time and labor efficient, practical instrument that can be used to assess the hand-related activity level for clinical and research purposes in patients with stroke.

Key Words: Cerebrovascular accident; Hand; Outcome assessment (health care); Rehabilitation.
performing office tasks, and other general items. The questions ask how much difficulty the person has in performing the tasks without the help of any assistive devices and/or personal assistance. It takes only a few minutes to complete the questionnaire, and no equipment or trained health care professionals are needed. The DHI shows promise as a reliable and valid assessment tool in patients with RA, osteoarthritis (OA), systemic sclerosis, and in those receiving hemodialysis. This questionnaire has not yet been used to assess hand-related activity limitations in patients with stroke. Therefore, the aim of this study was to examine the reliability, validity, and responsiveness of the DHI in the assessment of hand-related activity limitation in patients with hemiplegia after stroke.

METHODS

Participants

The study included 56 consecutive inpatients with hemiplegia after stroke, whom we recruited from 3 rehabilitation centers. Their mean ages and time since stroke ± standard deviation (SD) were 62.0 ± 14.2 years and 84.0 ± 52.6 days, respectively. The inclusion criteria were: having hemiplegia due to stroke, no previous pathology of the arm and/or hand, and cognitive ability to understand the questionnaire. Motor recovery level of the upper extremity was not an inclusion or an exclusion criterion. All participants received a conventional stroke rehabilitation program, 5 days a week, for 4 weeks. The conventional program is patient-specific and consists of neurodevelopmental facilitation techniques, physiotherapy, occupational therapy, and speech therapy (if needed). The baseline characteristics of the participants are summarized in Table 1. Approval was obtained from the Medical Ethics Committee prior to the study.

Sample Size

We determined the required sample size by using Number Cruncher Statistical Systems. Using an estimate of .90 for the lower-bound confidence interval (CI) (as per findings in patients with rheumatic conditions), a sample size of 56 participants with 2 observations per participant would achieve 84% power to detect an intraclass correlation coefficient (ICC) of .95.

Measures

We assessed body functions and structures of the affected arm and hand of the patients in terms of motor recovery (Brunnstrom stages), spasticity (Modified Ashworth Scale [MAS]), and sensation (light touch and joint position sense).

The DHI and FIM self-care items were used to measure activity limitation of the hand.

Brunnstrom Stages

Brunnstrom defined 6 sequential stages of motor recovery and described how the hemiplegic arm, hand, and leg progress through these stages as an acceptable method for assessing recovery in patients with hemiplegia. The 6 stages of the Brunnstrom for the hand are: (1) flaccidity; (2) little or no active finger flexion; (3) mass grasp, use of hook grasp but no release, no voluntary finger extension, and possibly reflex extension of digits; (4) lateral prehension, release by thumb movement, semi-voluntary finger extension, with small range, (5) palmar prehension, possibly cylindrical and spherical grasp, awkwardly performed and with limited functional use, voluntary mass extension of digits, with variable range, and (6) all prehensile types under control, skills improving, full-range voluntary extension of digits; individual finger movements present, but less accurate than on the opposite side.

Modified Ashworth Scale

We used the MAS to grade spasticity for arm and hand. The MAS is a 5-point ordinal rating scale with good interrater reliability designed to measure muscle tone. Higher MAS scores indicate worse spasticity.

Sensory Evaluation

Light touch and joint position sense of the paretic hand were recorded. We assessed light touch sense using a piece of cotton wool after the patient was familiarized to normal sensation on the anterior chest and forehead. The results were then categorized as normal, impaired, and absent. We assessed joint position sense by holding the patient’s index finger laterally at the distal interphalangeal joint and moving it up or down a few millimeters. The patient was asked to report the direction of movement relative to the last position and this was graded as normal, abnormal, or absent.

FIM Instrument

The FIM instrument is the functional status component of the Uniform Data System for Medical Rehabilitation. It is widely used in rehabilitation centers and has properties useful for stroke investigators. It contains 18 items that measure independent performance in self-care, sphincter control, transfers, locomotion, communication, and social cognition. FIM scores range from 1 to 7: a FIM item score of 7 is categorized as “complete independence,” whereas a score of 1 is “complete dependence” (performs <25% of task). Scores below 6 require another person for supervision or assistance. The FIM self-care items were used in the present study. The reliability and validity of the Turkish version of the FIM has been well documented. The second author (GY) has been trained in administration of the FIM and certified. The other assessors (NS, KS, PB) were trained by the second author. All assessors had at least 2 years of experience with the FIM in patients with stroke.

Duruoz Hand Index

The DHI was developed as a self-report questionnaire that can be routinely used to assess hand-related activity limitation in patients with RA. It has been cross-validated for outcome assessment of hand-related activity in patients with OA, systemic sclerosis, and those receiving hemodialysis. Its

Table 1: Baseline Characteristics of the Study Population (N=56)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y)</td>
<td>62.0 ± 14.2</td>
</tr>
<tr>
<td>Sex (women/men)</td>
<td>23/33</td>
</tr>
<tr>
<td>Education</td>
<td></td>
</tr>
<tr>
<td>Illiterate</td>
<td>9</td>
</tr>
<tr>
<td>Elementary school</td>
<td>36</td>
</tr>
<tr>
<td>High school, university</td>
<td>11</td>
</tr>
<tr>
<td>Type of lesion (ischemia/hemorrhage)</td>
<td>46/10</td>
</tr>
<tr>
<td>Paretic side (right/left)</td>
<td>23/33</td>
</tr>
<tr>
<td>Time since stroke (d)</td>
<td>84.0 ± 52.6</td>
</tr>
<tr>
<td>Brunnstrom stages of arm</td>
<td>2.7 ± 1.5</td>
</tr>
<tr>
<td>Brunnstrom stages of hand</td>
<td>2.4 ± 1.6</td>
</tr>
<tr>
<td>MAS of hand</td>
<td>1.5 ± 0.9</td>
</tr>
<tr>
<td>MAS of elbow</td>
<td>1.5 ± 0.9</td>
</tr>
</tbody>
</table>

NOTE: Values are n or mean ± SD.
Abbreviation: MAS, Modified Barthel Index.

reliability has been shown in patients with RA, OA, and systemic sclerosis. Its responsiveness has also been shown during the course of the disease and after hand surgery in patients with RA.

It contains 18 items on hand ability in the kitchen, during dressing, while performing personal hygiene, office tasks, and other general items (appendix 1). In previous studies, 3 factors were extracted for the DHI. The first factor represents activities requiring force and rotation (questions 2, 3, 5, 6, 11, 12, 15, 18), the second factor represents activities requiring dexterity and precision (questions 1, 4, 7–10), and the third factor represents dynamic activities requiring flexibility of the first 3 fingers. Scores for kitchen tasks range from 0 to 40. Scores for dressing, hygiene, and office tasks range from 0 to 10. Scores for the “other” category range from 0 to 20. Persons rate their ability from 0 (no difficulty) to 5 (impossible to do). The questionnaire yields a total score from 0 to 90, takes about 3 minutes to complete, and 6 levels of answers allow a more sensitive grading of hand-related activity limitation. A higher score indicates greater activity limitation or more difficulty.

Design

We assessed the patients on 3 different occasions. Two baseline measurements were performed with a 24-hour interval in between; the study population was considered to be clinically stable during the period between these 2 baseline measurements. A third follow-up assessment measurement was made immediately after completing a 4-week inpatient rehabilitation program. Figure 1 shows the flow of patients from recruitment to completion of the study. Because we aimed to measure activity limitation, we asked the patient to assess their ability without any hand preference. Literate patients completed the questionnaires by themselves and interviewers read the questions to illiterate patients.

Data Analysis

We analyzed the data using SPSS for Windows. Our data showed normal distribution. We calculated the means and SDs for all measures. Reliability refers to the reproducibility (the degree to which the score is free from random error) and the internal consistency of the instrument. In the present study, test-retest reliability and internal consistency of the DHI was measured by ICCs and the Cronbach α coefficient, respectively. Test-retest reliability measures the stability over time by administering the same test to the same subjects at 2 points in time; it is commonly evaluated using correlation statistics such as ICC, Pearson or Spearman coefficients, and κ coefficients. Internal consistency assesses the homogeneity of the scale items. It is generally examined using Cronbach α statistics or split-half reliability. Item-to-item and item-to-scale correlations are also accepted methods.

We calculated the minimal detectable change (MDC) between first and third assessment of the DHI. The MDC is a clinically relevant measure suggesting the change that might be expected because of an intervention rather than sampling error at a significance level of .05. The MDC is the 95% CI of the standard error (SE) of measurement multiplied by the square root of 2; that is, $MDC = 1.96 \times \sqrt{2 \times SE}$ of measurement. From the error variance, the SE of measurement was calculated as its square root. The SE of measurement is the SD of the population of all possible measurement errors. The SE of measurement is an estimate of how much a score is likely to vary with repeated measurements.

Validity is a term for how well an instrument measures what it intends to measure. Forms of validity include face, content, construct, and criterion. Concurrent, convergent, discriminative, and predictive validity are all considered to be forms of criterion validity.

The Spearman correlation coefficient was used to show the association between DHI and FIM self-care items (construct validity). Discriminative power of the DHI was investigated by comparing the data of patients with dominant side paresis versus nondominant side paresis, using an independent sample $t$ test.

Responsiveness is the sensitivity to change within a patient over time, which might be indicative of therapeutic effects. The usefulness of an outcome measurement is related to its sensitivity to change (responsiveness). Responsiveness is most commonly evaluated through correlation with other change scores, effect sizes, standardized response means, relative efficiency, sensitivity and specificity of change scores, and receiver operating characteristics analysis. We used effect size and paired $t$ test to assess responsiveness. The effect size is defined as the mean change in scores between the baseline and the follow-up visit divided by the SD of the baseline score. A higher effect size indicates greater responsiveness. A negative value indicates that the mean score at the baseline visit is smaller than the mean score at the follow-up visit. A paired $t$ test was used to determine the statistical significance of change scores.

RESULTS

Of the 100 consecutive stroke survivors that could have been recruited, 85 who met the inclusion criteria were enrolled in the study and had a baseline measurement; the second assessment was performed on 70 patients, and the third assessment was performed on 56 patients. Statistical analysis was done on these 56 patients (see Fig 1).

Light touch and/or joint position sensation was impaired in 16 patients (30%). Table 2 presents data for the baseline performance of stroke patients on the DHI and the FIM self-care items. The DHI and FIM self-care scores suggested moderate to severe activity limitation. The SE of mea-
Abbreviations: LE, lower extremity; UE, upper extremity.

NOTE. Values are mean ± SD. Abbreviations: LE, lower extremity; UE, upper extremity.

Reliability

We determined the test-retest reliability of the DHI between the first and second assessments using the ICC. The accepted evaluation criteria and standards for ICC values are as follows: values of .75 or greater represent excellent reliability; values between .40 and .74 represent adequate reliability; values of .40 or lower represent poor reliability. In the present study, we investigated the test-retest reliability and internal consistency for the DHI on both the individual section and total scores (Table 3). Each section showed ICCs of .99. We assessed internal consistency of the DHI with the Cronbach α; it is suggested that the α should be .80 and over for excellent internal consistency. In our study, Cronbach α coefficients for the DHI sections ranged from .91 to .97, which shows excellent homogeneity of the items in the questionnaire (see Table 3).

Validity

We assessed construct validity by Spearman correlation coefficients (Table 4). Correlation coefficients of .60 or greater represent excellent correlation, values between .31 and .59 represent adequate correlation, and values of .30 or lower represent poor correlation. In the present study the correlation between the DHI and the FIM self-care items was excellent. An adequate correlation was found between the DHI and Brunnstrom stages, and between the sensation of the affected upper extremity. There was no correlation between the DHI and spasticity level of the affected upper extremity. Discriminant validity was shown between subgroups of participants defined according to the affected side; as expected, patients with nondominant side paresis showed a better performance on the DHI, and the difference between the 2 groups was statistically significant.

Responsiveness

We assessed responsiveness of the DHI by calculation of the effect size and paired t test. There was a significant improvement in the DHI score after 4 weeks of inpatient rehabilitation program. The mean ± SD paired difference value was 6.5 ± 18.65 (P < .05). The effect size was .24.

DISCUSSION

Interventions to improve the activity level of stroke patients are often complicated and they require valid and reliable outcome measurements. The findings of this study revealed the reliability, construct validity, and responsiveness of the DHI, and provided support for its use as an instrument to measure hand-related activity limitation in stroke patients. Previous studies have shown that the DHI has excellent reliability in patients with systemic sclerosis, RA, and OA. In this study, we investigated the test-retest reliability and internal consistency of all individual sections. We found excellent test-retest reliability and internal consistency for the DHI on both the individual section and on the total scores for patients with stroke.

One of the most common distribution-based change indexes is the MDC, which is also known as smallest detectable difference or reliability change index. The MDC represents the minimal amount of change that is unlikely to be due to chance variation in measurement. In this study, the MDC of the DHI was 1.4 points, which means that any change between the 2 assessments exceeding 1.4 points in value can be attributed to true change but not to random measurement error or to change variation. From an examination of the index, it is unlikely that a change of less than 1.4 points would be considered to be clinically important; therefore, we conclude that the DHI would detect a clinically important change. Haley and Fragala-Pinkham suggested that journals can encourage the reporting of MDC values with the same regularity as clinical significance or effect sizes. They noted that the interpretation of clinical significance would become transparent and more commonly accepted if we were informed about MDC.

Adequate validity has been established for the DHI in patients undergoing hemodialysis, and in those with systemic sclerosis, RA, and OA. To estimate the construct validity of the DHI scores in stroke patients, we investigated the association between the DHI scores and the scores of the body function assessments (motor recovery, sensation, spasticity) and the FIM self-care items, using Spearman correlation coefficients. The DHI showed an excellent correlation with the FIM self-care items. This was expected, because both questionnaires evaluate limitation at the level of activities in daily life. An adequate correlation was found between sensation and the DHI and Brunnstrom stages. Because these stages assess body functions and structures (whereas DHI assesses activity level) a correlation of this magnitude would be acceptable. We did not find correlation between the DHI and spasticity level.
Previous studies\textsuperscript{49,50} suggested that function is associated more with motor control than muscle tone. The majority of the patients in our study had low spasticity; hence there was low variability which would limit the size of correlation.

We determined the discriminative power of the DHI by comparing the scores of patients with dominant side paresis versus nondominant side paresis, using independent sample t tests. In our study, the DHI successfully discriminated between stroke patients with dominant side paresis and those with nondominant side paresis. It has previously been reported that the DHI is a discriminative functional test for assessing activity limitation in patients undergoing hemodialysis and those with OA.\textsuperscript{1,11,30}

In this study, we found the DHI to be responsive to changes in the clinical characteristics of stroke patients after a 4-week inpatient rehabilitation program (effect size, .24; P<.05). Previous studies reported similar effect sizes of the DHI for patients with RA\textsuperscript{41} and hand OA.\textsuperscript{30} Our study was not designed to compare the responsiveness of the DHI with other activity measures. It may be that other activity measures are more responsive; however, this needs to be determined and weighed against the time, equipment, and training that they require.

Study Limitations

Stroke patients recruited in this study were referred from all over the country for inpatient rehabilitation. Generally, an estimated 50% of the stroke population is referred to a rehabilitation center if they cannot return home directly after dismissal from the hospital. A potential limitation of this study is the generalizability of the results. According to our inclusion criteria, our findings and conclusions are based on the population of subacute stroke inpatients without severe cognitive deficits but with severe motor impairment of hand and upper extremity. Most severe strokes with cognitive problems and least severe outpatient strokes were eliminated at the beginning of the study. Further studies evaluating chronic stroke outpatients with less motor impairment should be performed.

CONCLUSIONS

The DHI is a time and labor efficient, practical, self-reported instrument that can be used to assess the hand-related activity level in patients with stroke. The advantage of the DHI in clinical practice may be its relative simplicity and the shorter time required for training and administration. Our study indicates that it can be used as both a descriptive tool and as an outcome measure.

APPENDIX 1: THE ENGLISH VERSION OF THE DURUOZ HAND INDEX

Answers to the questions:

<table>
<thead>
<tr>
<th>1-Can you hold a bowl?</th>
<th>2-Can you take a full bottle and raise it?</th>
<th>3-Can you hold a plate full of food?</th>
<th>4-Can you pour liquid from a bottle into a glass?</th>
<th>5-Can you unscrew the lid from a jar opened before?</th>
<th>6-Can you cut meat with a knife?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes without difficulty</td>
<td>Yes, with a little difficulty</td>
<td>Yes, with some difficulty</td>
<td>Yes, with much difficulty</td>
<td>Nearly impossible to do</td>
<td>Impossible</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

Answer the following questions regarding your ability without the help of any assistive devices.

In the kitchen:
1. Can you hold a bowl?
2. Can you take a full bottle and raise it?
3. Can you hold a plate full of food?
4. Can you pour liquid from a bottle into a glass?
5. Can you unscrew the lid from a jar opened before?
6. Can you cut meat with a knife?
7. Can you prick things well with a fork?
8. Can you peel fruit?

Dressing:
9. Can you button your shirt?
10. Can you open and close a zipper?

Hygiene:
11. Can you squeeze a new tube of toothpaste?
12. Can you hold a toothbrush efficiently?

In the Office:
13. Can you write a short sentence with a pencil or an ordinary pen?
14. Can you write a letter with a pencil or an ordinary pen?
15. Can you turn a round door knob?
16. Can you cut a piece of paper with scissors?
17. Can you pick up coins from a table top?
18. Can you turn a key in a lock?  

Total Score:

References

Aerobic Capacity After Traumatic Brain Injury: Comparison With a Nondisabled Cohort

Kurt A. Mossberg, PhD, PT, Danielle Ayala, MPT, Tracey Baker, MPT, Justin Heard, MPT, Brent Masel, MD


Objective: To compare aerobic capacity of people recovering from traumatic brain injury (TBI) with an age- and sex-matched group of nondisabled sedentary people.

Design: Descriptive comparative study of peak and submaximal physiologic responses.

Setting: Residential postacute treatment center.

Participants: Convenience sample of 13 people with TBI and 13 age- and sex-matched nondisabled subjects. All subjects could walk 5.3kph (3.3mph), follow 2-step commands, and comply with testing using the gas collection apparatus.

Interventions: Not applicable.

Main Outcome Measures: Subjects performed a graded maximal treadmill test during which heart rate, minute ventilation (Ve), oxygen consumption (Vo2), carbon dioxide production, and respiratory exchange ratio (RER) were measured every minute until exhaustion. Ventilatory equivalents for oxygen (Ve/Vo2) and oxygen pulse were calculated.

Results: Subjects recovering from TBI had significantly lower peak responses for heart rate, Vo2, Ve, and oxygen pulse TBI (P.<.01). Peak RER and Ve/Vo2 were similar. There were significant differences in submaximal responses for Ve/Vo2 and oxygen pulse.

Conclusions: Patients with TBI were significantly more deconditioned than a comparable group of nondisabled sedentary people. Participation in cardiorespiratory fitness programs after TBI should be encouraged to prevent secondary disability.

Key Words: Brain injuries; Head injuries, closed; Physical endurance; Rehabilitation.

© 2007 by the American Congress of Rehabilitation Medicine and the American Academy of Physical Medicine and Rehabilitation

An estimated 1.4 million people sustain a traumatic brain injury (TBI) in the United States every year. A large percentage of these people are relatively young, with mild brain injuries that result in few or no physical impairments. Because they are young, they can expect to live for many years with the potential of developing age-related chronic disabilities. Many such disabilities are associated with physical inactivity and a sedentary lifestyle. Some negative results of inactivity are poor stamina, reduced muscle strength, and limited flexibility. It is well established that generally, people who live sedentary lifestyles are at greater risk for coronary heart disease, hypertension, thromboses, osteoporosis, obesity, certain cancers, and non-insulin-dependent diabetes mellitus.

Presumably the same risks faced by the general population exist for people recovering from TBI. The combination of living with a disability and being sedentary increases the risk of developing secondary conditions. Unfortunately, longitudinal studies that describe chronic disease development and its relation to physical activity levels in TBI patients have not been reported. The effects of these health problems are confounded in people with disabilities. Jankowski and Sullivan provided data that strongly suggest that peak aerobic capacity is related to employment productivity in people recovering from TBI; they have a diminished tolerance for continuous physical activity and chronic fatigue is a common complaint, even years after injury. For these reasons, it is crucial that they become as active as is feasible.

The degree of aerobic or endurance capacity limitation in recovering TBI patients is not well documented. It has been estimated that their peak aerobic capacities are from 65% to 74% of normative values. The certainty of these estimates is questionable for several reasons. First, many patients with TBI have physical impairments. Becker et al did a direct comparison of nondisabled sedentary subjects and patients with TBI but 58% (11/19) of the patients had residual motor impairments. In addition, that study assessed submaximal responses but not peak responses. Other factors that lead to questions about these estimates relate to the testing protocols. It is not known whether testing protocols used with the subjects with TBI and the normative values from which the estimates were derived were identical. The above estimates are from studies performed with the treadmill as the testing modality, while others utilized the cycle ergometer. Bhambhani et al tested TBI patients on the cycle ergometer and compared their results with those of people without health problems who performed a treadmill protocol. Investigators have compared cycling with treadmill ambulation in the nondisabled and after TBI and found peak oxygen consumption (Vo2peak) was significantly higher on a treadmill than on a bicycle ergometer. Furthermore, it is well known that Vo2peak can vary depending on the treadmill protocol. Consequently, these estimates are based on comparisons that may not be valid.

We know of no data that directly compares the aerobic capacity of recovering TBI subjects with few, if any, physical impairments with the aerobic capacity of nondisabled subjects of similar age and sex when performing identical graded exercise tests. Such a comparison is necessary in order to draw accurate conclusions about the cardiorespiratory fitness of patients recovering from TBI. Our purpose in this study was to make just such a comparison.


**Table 1: Age Characteristics of Study Subjects**

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Mean ± SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBI (n=13)</td>
<td>30.4±8.1</td>
<td>19–47</td>
</tr>
<tr>
<td>Age at injury (y)</td>
<td>31.2±7.9</td>
<td>22–48</td>
</tr>
<tr>
<td>Nondisabled (n=13)</td>
<td>31.7±6.8</td>
<td>23–49</td>
</tr>
</tbody>
</table>

Abbreviation: SD, standard deviation.

**METHODS**

**Participants**

The 13 participants with TBI were residents in a residential postacute treatment center and had been admitted primarily for neuropsychologic and vocational rehabilitation. To be included in the study, they had to be free of overt cardiovascular disease, could not be taking cardiovascular medications, could follow 2-step commands, could comply with the gas collection apparatus, and could walk independently at a treadmill speed of 5.3kph (3.3mph). Patients were screened and clinical examinations found no obvious balance and musculoskeletal impairments. Cognitive screening showed that 9 subjects scored average on attention, language comprehension, sentence repetition, and naming of simple objects. The greatest cognitive impairments in the TBI group were in the areas of verbal and auditory memory, visual memory, and visual constructional skills. Table 1 shows the age characteristics of all 26 subjects. Patients with TBI were tested an average of 10.4±9.5 months (range, 6wk to 32mo) after injury. Initial examination postinjury showed 11 subjects had severe injuries (Glasgow Coma Scale [GCS] score range, 3–8), 1 subject had a moderate injury (GCS score, 9), and 1 subject had a mild injury (GCS score, 13). The injuries resulted from motor vehicle collisions (n=10), falls (n=2), and an assault (n=1). An equal number of age- and sex-matched nondisabled subjects were studied after a general screening for cardiovascular health. The apparently healthy people were recruited through personal contact with one of the investigators. They were faculty members, students, and staff members of a local university and the residential treatment center. Twelve men and 1 woman in each group gave their written informed consent to participate. Institutional review boards at both the residential treatment center and the university approved all procedures. All subjects were familiar and comfortable with treadmill ambulation and none were participants in regular formal aerobic exercise programs prior to testing. Five patients with TBI and 1 nondisabled subject were smokers.

**Instrumentation**

Cardiorespiratory measures were evaluated using an automated metabolic cart and resting and exercise heart rates were monitored through electrocardiography. The metabolic cart was calibrated before and immediately after each test, using gases of known concentrations. The pneumotach was calibrated with a 3-L syringe.

We used a standard treadmill that was equipped with side railings and handles at the front. Subjects were told to use the hand supports for balance only and were constantly reminded during the test to use them for that purpose only. We used a waist belt that was free of attachments to the treadmill to secure electrocardiography lead wires and to facilitate manual assistance if a subject’s balance was compromised. At no time during the testing, however, were the subjects given manual assistance.

**Procedure**

All subjects performed a graded maximal treadmill test that was a modification of the Balke-Ware protocol. There was a 2-minute warm-up at 1% incline during which the treadmill speed was gradually increased to a velocity of 5.3kph (3.3mph). During the test, the velocity was kept constant while the incline was increased 2% each minute. This modified protocol has been used previously in people recovering from TBI and has been shown to be reliable in this population.

Heart rate, electrocardiogram, oxygen consumption (VO2), carbon dioxide production (VCO2), minute ventilation (VE), and respiratory exchange ratio (RER) were monitored continuously throughout the test. Subjects were encouraged to give their best effort. The test was stopped if the subject asked to stop, if continued ambulation became unsafe, or when 2 of the following 3 peak criteria were achieved: (1) the subject’s V02 reached a plateau with an increase in workload, (2) the heart rate reached 90% of age-predicted maximum (220 – age), and (3) the subject achieved an RER equal to or greater than 1.15. Peak effort was defined as the highest values recorded during the last 30 seconds of exercise.

**Data Analyses**

Oxygen pulse, a noninvasive estimate of cardiac stroke volume, was calculated as the ratio of oxygen consumed (in L/min) to heart rate (in bpm). Ventilatory equivalent for oxygen (Ve/Vo2) is a measure of the ventilatory muscle effort required to exchange a given amount of oxygen. It was calculated as the ratio of minute ventilation (in L/min) to oxygen consumed (in L/min).

We analyzed peak data for heart rate, V02, oxygen pulse, Ve, Ve/Vo2, and RER with paired samples t tests. Because we studied multiple variables, a Bonferroni adjustment was made in order to maintain an α level of .05. We used a multivariate repeated-measures 2-way analysis of variance to test for the main effects of time and group as well as time-by-group interaction for the submaximal responses up to and including minute 6. Data beyond minute 6 were excluded from this submaximal analysis because the first subjects in the TBI group started to reach their peaks after minute 7. Pearson product moment correlations on V02,peak and heart rate versus subject demographics were performed. All analyses were carried out at the α level of .05. Statistical analyses were performed using a personal computer–based statistical software program.

**RESULTS**

Table 2 shows a comparison of average peak values for heart rate, V02, Ve, oxygen pulse, Ve/Vo2, and RER between the 2 subject groups. After the Bonferroni adjustment all variables except Ve/Vo2 and RER differed statistically (P<.01). The fact that RER was similar in the groups suggests that the level of effort (level of anaerobic metabolism) was similar. Patients...
with TBI were more likely to reach an oxygen plateau and an RER equal to or greater than 1.15 and less likely to reach 90% of their age-predicted maximal heart rate.

Figure 1 illustrates the differences in the distribution of VO2peak values (in mL·kg⁻¹·min⁻¹) and oxygen pulse (in mL/beat) between participants with no disability and those recovering from TBI. The greater oxygen pulse for the subjects without disability suggests that cardiac stroke volume was lower in patients with TBI. Correlations between the initial severity of injury (GCS) and VO2peak or heart rate (table 3) were not significant. There was a modest positive correlation (r=0.45) between the time since injury and the peak aerobic capacity but this was not significant (P=0.12), and recovery time only predicted 20% of the variance in VO2peak for the subjects with TBI. There was a significant inverse relation (P<0.05) between the age of a patient and his/her peak aerobic capacity. There was a much more modest and insignificant relationship (r=−0.36) between age and VO2peak for the nondisabled subjects.

Examination of submaximal responses revealed that subjects with TBI consumed slightly greater amounts of oxygen during the initial exercise stages (fig 2). During the first 6 minutes of submaximal exercise, however, there was no significant group effect (P=0.61). During the first 6 minutes of submaximal exercise, subjects without disability were noted to begin consuming greater amounts of oxygen consistent with the greater workload and their greater peak aerobic capacity. Similarly, the TBI subjects had slightly greater submaximal minute ventilation than the nondisabled cohort but there was no significant group effect (P=0.9, data not shown). Figure 3 illustrates the submaximal ventilatory equivalents for oxygen for both groups through minute 6. There were significant differences between the nondisabled and TBI groups (group effect, P=0.002). Subsequent pairwise comparisons revealed significant differences after the second minute of exercise. In addition, there was a significant interaction between group and time of exercise (P=0.014).

Subjects with TBI tended to have higher submaximal heart rates than nondisabled subjects in the first 6 minutes of walking but there was no significant group effect (P=0.16, data not shown). Only until late in the test did the nondisabled subjects start to attain higher heart rates consistent with their higher peak heart rate (see table 2). Figure 4 illustrates the submaximal oxygen pulse for the 2 groups. Again, there were significant differences between the 2 groups after minute 2 (group effect, P=0.006).

**DISCUSSION**

Before this study, there had been no direct comparisons of cardiopulmonary fitness of a healthy population with that of people recovering from TBI who had few, if any, physical impairments. We found significant differences in several variables related to cardiorespiratory fitness. Our results indicate that patients recovering from TBI have cardiovascular and pulmonary limitations in endurance activities. One of the more surprising differences between the groups was the peak heart rate response. We found that all nondisabled subjects attained greater than 90% of age-predicted maximum heart rate but only 8 of 13 patients recovering from TBI achieved this level. This fact, combined with the lower oxygen pulses (see figs 1, 4), suggests a limitation in exercise cardiac output.

We also found pulmonary limitations to peak endurance activity. A study by Becker et al found the submaximal $\dot{V}O_2$ to be 30% higher in people with TBI than in nondisabled subjects. We found differences on the same order for most minutes of submaximal exercise. Peak values for the 2 groups in our study, however, did not differ due to the relative differences in VO2peak and VE (76% of TBI and 68% of normative values, respectively). The higher submaximal
VE/VO₂ (see fig 3) indicates that breathing is less efficient and people with TBI must breathe harder during physical activity to exchange a given volume of oxygen. Becker’s group also found oxygen pulse to be approximately 35% lower in patients with TBI than in healthy subjects.22 In those subjects, peak aerobic capacity was less than or equal to the 5th percentile when compared with the age- and sex-matched comparison group.21 Comparisons should be made with caution.

Balke and Ware16 studied the work capacity of male Air Force personnel and rated it as a function of VO₂peak and test duration. According to their study, subjects were rated “very poor” if their oxygen uptake was 25 to 30mL·kg⁻¹·min⁻¹. A VO₂peak of less than 25mL·kg⁻¹·min⁻¹ was considered an “inferior” level of fitness. In our sample, 9 of 12 men with TBI fell into 1 or the other of these 2 “most unfit” classifications. The average for the nondisabled men (35.3mL·kg⁻¹·min⁻¹) would be rated as borderline “fair,” a rating given to people whose oxygen uptake is 35 to 40mL·kg⁻¹·min⁻¹.16 Because our protocol is a modification of the Balke-Ware protocol, we assume that comparisons are reasonably valid. We have, however, modified the protocol; speed is gradually increased to 3.3mph in the first 2 minutes and the incline increases 2% every minute instead of the 1% used by Balke and Ware.

In another study of healthy subjects,22 untrained men had VO₂peak that approximated 44mL·kg⁻¹·min⁻¹ and untrained women had VO₂peak of approximately 39mL·kg⁻¹·min⁻¹. For our subjects without disability, on average the 12 men and 1 woman consumed 20% and 7% less oxygen at peak work, respectively. The 12 men and 1 woman with TBI consumed 39% and 28% less oxygen at peak work, respectively. Because the details of the testing protocol used in the untrained subjects were not provided by Saltin and Astrand,21 comparisons should be made with caution.

Using a different data set for normative values, peak aerobic capacity for all patients with TBI was less than or equal to the 5th percentile when compared with the age- and sex-matched healthy subjects.22 In those subjects, peak aerobic capacity ranged from the 5th to the 75th percentile; the median was the 10th percentile. Even greater caution should be taken if these normative values are used for comparison because peak aerobic capacity is estimated based on heart rate responses to submaximal “stepping” activity. From Pollock and Wilmore22 and the other previously cited studies cited confirm that our nondisabled cohort was comprised of sedentary people not undergoing endurance training. Moreover, it further supports our finding that our subjects with TBI had an aerobic capacity that was 76% of the healthy age- and sex-matched comparison group when an identical testing protocol was used. Our results are supported by other indirect comparisons in which patients with TBI approached only 67%6 and 74%9 of age-predicted maximal VO₂ during treadmill testing and 65% during cycle ergometry.20

There are probably many factors that contribute to the lower physical work capacity in patients recovering from TBI that have few, if any, physical impairments. The lack of activity is most likely the major contributor. This can result from a lack of motivation and other cognitive deficits, and from a lack of effective education and attention given by rehabilitation professionals to the issue of cardiorespiratory fitness as a key component of both physical health and mental health. Other factors that probably play a role are the indirect effects of certain medications and subtle balance impairments that make unsupervised high intensity training unsafe. More formal investigative work is necessary to know with certainty the reasons for decreased cardiorespiratory fitness.

We found a significant inverse relation between age and peak aerobic capacity in the patients recovering from TBI (see table 3). In the nondisabled subjects, however, the age relationship is much weaker, probably because of the lower variability in VO₂peak (see fig 1). The data for patients recovering from TBI only weakly suggest that the longer the time since injury, the greater the peak aerobic capacity. This was somewhat surprising, but still encouraging, in that it supports the notion that endurance levels can increase over time with a resumption of physical activity. Because of the significant differences related to both submaximal and peak cardiorespiratory fitness, however, patients should be encouraged to engage in more formal, properly prescribed endurance training programs.

Decreased cardiorespiratory fitness leads to increased fatigability, thus hindering one’s ability to perform a work task. Fatigue is a common complaint among people who have a TBI,7,8 and endurance training has the potential to delay the onset of fatigue. Therefore, aerobic training should be a component of all rehabilitation programs so that TBI patients can increase their work levels during various endurance tasks and hasten their rehabilitation. Data have been published that suggest that aerobic training reduces fatigability and may be essential in the vocational rehabilitation and placement of people with TBI.6 Endurance training programs may play an important role in preparing TBI patients to better handle the sustained work demands occasioned by being employed. Physical conditioning programs have been advocated for many years for patients recovering from TBI.23 Studies that examined changes in physi-
AEROBIC CAPACITY AFTER TBI, Mossberg 319

Aerobic fitness is severely limited in patients recovering from TBI who have few, if any, other physical impairments. Our results point to limitations at both cardiac and pulmonary organ system levels. More than likely there is also a limitation in oxygen extraction by the active skeletal muscle, but this needs further study. Patients recovering from the sequelae of TBI who have no musculoskeletal impairments that preclude participation in cardiorespiratory training programs should be encouraged to do so. Longitudinal studies are necessary to determine if these training activities will mitigate the chronic fatigue experienced by a majority of these patients and prevent secondary disability resulting from a sedentary lifestyle.

Acknowledgments: We thank Charlie Milton, MS, PT, for his support in recruiting subjects and Julie Norcross for her expert technical assistance.

References


**Suppliers**


b. Cardio Perfect Inc, 1870 The Exchange, Ste 150, Atlanta, GA 30339.

c. Version 14.0; SPSS Inc, 233 S Wacker Dr, 11th Fl, Chicago, IL 60606.
The Development and Validity of the Salford Gait Tool: An Observation-Based Clinical Gait Assessment Tool

Brigitte Toro, PhD, Christopher J. Nester, PhD, Pauline C. Farren, PhD


OBJECTIVES: To develop the construct, content, and criterion validity of the Salford Gait Tool (SF-GT) and to evaluate agreement between gait observations using the SF-GT and kinematic gait data.

DESIGN: Tool development and comparative evaluation.

SETTING: University in the United Kingdom.

PARTICIPANTS: For designing construct and content validity, convenience samples of 10 children with hemiplegic, diplegic, and quadriplegic cerebral palsy (CP) and 152 physical therapy students and 4 physical therapists were recruited. For developing criterion validity, kinematic gait data of 13 gait clusters containing 56 children with hemiplegic, diplegic, and quadriplegic CP and 11 neurologically intact children was used. For clinical evaluation, a convenience sample of 23 pediatric physical therapists participated.

INTERVENTIONS: We developed a sagittal plane observational gait assessment tool through a series of design, test, and redesign iterations. The tool’s grading system was calibrated using kinematic gait data of 13 gait clusters and was evaluated by comparing the agreement of gait observations using the SF-GT with kinematic gait data.

MAIN OUTCOME MEASURES: Criterion standard kinematic gait data.

RESULTS: There was 58% mean agreement based on grading categories and 80% mean agreement based on degree estimations evaluated with the least significant difference method.

CONCLUSIONS: The new SF-GT has good concurrent criterion validity.

KEY WORDS: Cerebral palsy; Gait; Physical therapy techniques; Rehabilitation; Reliability and validity.

© 2007 by the American Congress of Rehabilitation Medicine and the American Academy of Physical Medicine and Rehabilitation

INSTRUMENTED GAIT ANALYSIS remains the criterion standard assessment tool for the management of gait abnormalities, although the interpretation of gait analysis data for use in making clinical decisions varies. In routine clinical practice, however, access to an instrumented gait laboratory is relatively rare, and has resulted in the development of a variety of observational gait assessment (OGA) tools as an alternative. The process by which these tools have been developed, however, is often unclear and there is scarce evidence to support their validity. In the context of gait analysis, validity refers to the degree to which the assessment measures the actual events of gait. Face validity is the lowest level of validity and is based on the observer’s personal opinion. Construct validity is determined by theoretical reasoning that a gait assessment tool adequately measures selected gait variables. When a gait assessment tool is believed to include the domains that are required to adequately assess gait, its content is considered valid (content validity). Criterion validity, the highest level of validity of a gait assessment tool, is evaluated by comparing the results obtained by the tool to the criterion standard measurement of gait, which is instrumented analysis of gait kinematics, kinetics, and muscle activity. The use of tools that have criterion validity is justified because we can have confidence in their ability to accurately reflect actual gait events.

The development of most existing OGA tools appears to have been based on their clinical construct validity, rather than by demonstration of their criterion validity through comparison with quantitative kinematic gait data. Those tools that have been compared with the criterion standard have only poor to moderate validity.

The Hugh Williamson Gait Laboratory Scale, a modified version of the Physician Rating Scale (PRS), had poor criterion validity when it was compared with the quantitative gait assessment of sagittal plane foot and knee joint kinematics of 25 children with cerebral palsy (CP) gait, using 4 experienced raters (κ range, .46–.61). The Observational Gait Scale (OGS), another variant of the PRS, also had modest validity when compared with quantitative kinematic data (κ=.69; range, .38–.94) in the assessment of 20 children with spastic diplegia by 2 experienced assessors. That investigation was limited to the scale’s first 4 sections (knee position at mid stance, initial foot contact, foot position at mid stance, timing of heel rise). There was poor agreement between 3-dimensional gait data and 2 experienced observers using the Visual Gait Assessment Scale (VGAS), the most recent variant of the PRS, in the assessment of the gait of 31 children with CP hemiplegia. The mean κ scores ranged from .05 to .51 (mean, .22) for 4 parameters (hip position in terminal stance, hip position in mid-swing, knee peak extension in terminal stance, knee peak flexion in swing). The validity of the Edinburgh Visual Gait Score (EVGS) was measured by reporting agreement between scores and quantitative kinematics for each of the 10 numeric gait items that measure movement at the ankle, knee, hip, and pelvis. Percentage agreement between kinematic data and 5 experienced observers who assessed sagittal gait images from 4 children with CP and 1 neurologically intact child ranged from 47% for maximum knee extension in stance, up to 83% for maximum ankle dorsiflexion in swing (mean agreement, 64%).

The lack of objective information about existing observation-based gait assessment tools, coupled with evidence of the unmet needs of clinicians, led us to develop the Salford Gait...
Tool (SF-GT), a new clinically orientated OGA tool for therapists who manage gait problems of children with CP. This article describes the development of the tool’s construct, content, and criterion validity, using a combination of clinical experience and quantitative kinematic data, and then reports the levels of agreement between gait assessments made with the SF-GT and quantitative kinematic data.

METHODS

Initial Design of the SF-GT

The initial structure of the SF-GT was based on our previous experience, a review of existing tools, and reviews of video images of 10 children (6 boys, 4 girls; mean age, 7y; age range, 5−10y) with hemiplegic (n=3), diplegic (n=3), and quadriplegic (n=4) CP gait. All 10 gave their written consent to participate in the research and ethics approval was granted. The initial tool was designed to assess sagittal plane hip, knee, and ankle angular positions at 6 specific events during the gait cycle (initial contact, end double support, mid stance, start double support, toe-off, mid swing). We selected 6 events as a compromise between using a large number of events that would enable a comprehensive assessment but would be time consuming, or using a small number of events that would be quicker to complete but might provide too little detail. Moreover, the gait events needed to be precisely and repeatedly identifiable visually.

We devised a 5-point category scoring system (2, 1, 0, −1, −2) to describe the positions of the hip, knee, and ankle at the 6 gait events (3 joints, 6 gait events = 18 total assessments). Each scoring category would correspond to a specific range of angular positions. Importantly, we felt that the range of angular positions defining each category must be small enough to be sensitive to differences between different gait styles and changes resulting from clinical interventions, but large enough to be identified by the naked eye when observing gait on a video screen.

The sum of the 6 category scores for each joint would then represent that joint’s function over the entire gait cycle and provide a qualitative description of the entire joint pathology.

It was our intention that the final definitions of the boundaries between 1 scoring category and the next, as well as the boundaries between the summed category scores, would be based on quantitative kinematic gait data that described different types of gait pathology. To enable us to obtain early user feedback on the tool’s overall design, however, the boundaries between categories were provisionally set based on gait literature and clinical observation.

Development of Construct and Content Validity

This initial version of the tool was then evaluated for its user friendliness, physical layout, wording, and construct and content validity by 9 successive focus groups over a 1-year period. Each group had an average of 19 physical therapy students per group (total N=152), and 1 group included 4 physical therapists who were specialists in the fields of gait assessment, pediatrics, neurology, and musculoskeletal physical therapy. There were several testing, redesign, and testing iterations as changes were made to the tool at each stage.

Development of Criterion Validity

In this development stage, we used quantitative kinematic gait data to adjust the upper and lower boundaries of the scoring categories for each of the 3 joints at the 6 gait phases of the SF-GT. Such boundaries in existing tools appeared to be defined based on clinical experience rather than on formal evaluation of the gait patterns the tools would subsequently be used to evaluate. We assumed that careful adjustment of these boundaries would improve the SF-GT’s validity, sensitivity, and specificity.

In previous work we defined 13 gait styles by using cluster analysis of kinematic gait data from 56 children with a mixture of CP types and 11 neurologically intact children. Details of the kinematic data collection protocol and statistical analysis have been reported. We adjusted the SF-GT so that the upper and lower boundaries of each scoring category and the boundaries of the summed category scores reflected the mean joint positions and standard deviations (SDs) of the kinematic data of the children with each of the 13 distinct gait types. The calibration procedure was completed for the hip, knee, and ankle at all 6 gait phases. It was stopped when the scoring of the kinematic data of all gait types resulted in an adequate numeric and qualitative description of the entire joint pathology that was different from the other gait types.

Clinical Evaluation

To evaluate the clinical validity of gait assessments made with the SF-GT, the gait of 13 children was assessed by 23 pediatric physical therapists. The visual assessments were then compared with the quantitative assessment of the children’s gait. The observers were recruited from 9 National Health Service Trusts from the Greater Manchester area in the United Kingdom and their clinical expertise ranged from junior (n=1), senior II (n=2), senior I (n=18), to superintendent (n=2). All observers gave their written consent before participating. The 13 children each represented 1 of the 13 gait styles previously defined with cluster analysis (9 boys, 4 girls; age range, 6−16y; mean age, 9.5y). Eleven children had hemiplegia (n=4), diplegia (n=6), or quadriplegia (n=1) CP and 2 children were neurologically intact. The video data were collected during the same laboratory visit at which the kinematic gait data were collected. We used that data, which were part of a larger clinical gait analysis database, to define the 13 gait styles. Parents and children had previously given their consent to use the images for research purposes and the research was approved by the appropriate ethics committees.

Before making the gait assessments, the 23 observers were given general training on gait, video assessment of gait, and the use of the SF-GT. Observers were each allocated a workstation with a DVD player and television screen and did a trial gait assessment. The observers then worked individually without discussion to assess 1 gait cycle of 1 leg of each of the 13 children, using the SF-GT. They were permitted to work at their own speed, to review the gait cycles as often as required, and there was no time limit. The assessments were completed in from 3 to 5 hours.

Data Analysis

We analyzed the level of agreement between clinicians’ gait assessments using the SF-GT and the criterion standard kinematic gait assessments in 2 ways. First, we compared the frequency of agreement between the scoring categories estimated by each observer with the scoring categories assigned by the kinematic data (eg, if 5° of knee flexion at initial contact by the kinematic data were category 1, there was agreement if the observer also rated the knee as category 1). We also analyzed the extent of agreement by counting the frequencies of observer category scores that deviated by 1, 2, or more categories from the categories assigned by the kinematic data. The results of this analysis would enable direct comparison with the results of other OGA tools.
NOTE. Values are mean degrees ± SD.

Abbreviations: E, extension; F, flexion; NA, not applicable.

A disagreement, however, between observer category scores with categories assigned by the kinematic data may not always reflect the true closeness (or difference) between the observer assessment and kinematic data. For example, if an observer estimated the hip flexion to be 46° (category 2) and the actual kinematic data were 45° (category 1), this would have resulted in a disagreement by 1 category, even though there would have been only 1° difference between the assessments. Conversely, if an observer estimated the hip flexion to be 45° (category 1) and the actual kinematic degrees were 16° (also category 1), this would have resulted in a perfect agreement, even though the estimations would have differed by 29°. Therefore, the observer derived angles in degrees (recorded to 1°) were compared with the joint positions derived from the kinematic gait data (recorded to 0.1°). For this analysis, we used the least significant difference (LSD) method. The LSD represents a statistically significant difference between observer and kinematic data and occurs because of interclinician variation. It considers the distance between all the observations and the target kinematic data and is derived by multiplying the SD between 2 assessments (observer and kinematic data) with the relevant t distribution from the Student t test (with 1 degree of freedom, 2-tailed t test, .05 level of significance). If the observer’s estimated degree value lies outside the LSD range, then the difference between observers and kinematic degrees is statistically significant and too large to be considered acceptable. We used Excel for the statistical analysis.

**RESULTS**

**Criterion Validity**

We set the final definitions of the boundaries between each scoring category at the end of the calibration procedure. An example of the calibration result for the knee at initial contact for the 13 gait styles is given in table 1. The final gait assessment tool is detailed in appendix 1, which also provides information on normative gait data and guidelines for use of the tool.

**Clinical Evaluation**

**Agreement between category scores.** Observers agreed on 2900 category scores (58%; range, 28%–75%) of a possible 5004 category scores with the kinematic gait data (table 2). Regarding the magnitude of disagreement, of the 2104 category scores that disagreed, 95% (1991 category scores) disagreed by 1 category, 4% (93 category scores) disagreed by 2 categories, 1% (19 category scores) differed by 3 categories, and 1 observer’s category score differed by 4 categories from the kinematic category score.

Best agreement (83%) was for the hip joint of cluster 1 (gait style is mild crouch gait) and the lowest agreement (22%) was for the ankle joint of cluster 3 (gait style is moderate crouch gait). The knee joints were assessed more accurately than the hip and ankle joints (mean agreement: hip, 58%; knee, 61%; ankle, 56%). Cluster 10 (weak plantarflexion gait) was the gait style with the best agreement (75%). Cluster 3 (moderate crouch gait) had the least agreement (28%).

**Agreement between degrees.** We computed the LSD on a total of 4446 pairs of observer and kinematic data for all gait events (n=6) of all joints (n=3) and clusters (n=13). The degree data from 2 observers were not recorded on the SF-GT and therefore were not available for analysis. The mean LSD was 16.25° (range, 2.31°–63.20°). This means that the mean observations differed by 16.25° from the kinematic data. An average of 80% (range, 68%–91%) of gait observations lay within the LSD range of the kinematic data at the 6 phases of gait for all joints and clusters (table 3). Best agreement (91%) was for the hip joint of cluster 4 (gait style is severe crouch) and the lowest agreement (68%) was for the hip joint of cluster 2 (gait style is mobile crouch). The mean agreement between joints was similar (mean agreement: hip, 80%; knee, 80%; ankle, 81%). The clusters that demonstrated best agreement (both 84%) were clusters 6 (gait style is moderate equinus/knee extension) and 13 (gait style is normal gait). The cluster with the least agreement was cluster 2 (gait style is mobile crouch). Regarding the magnitude of agreement between the observers and the kinematic data, 86% of observations were within 25° of kinematic data, 79% of observations were within 20° of kinematic data, 62% of observations were within 15° of kinematic data, 50% of observations were within 10° of the kinematic data, and 33% of observations were within 5° of the kinematic data.
data, 24% within 10°, and 2.6% within 5°. Table 4 shows the data from 1 observer for mild equinus gait (cluster 5); this gait cluster was representative of a mean agreement of 80% between all observers.

**DISCUSSION**

There was 58% mean agreement between clinicians and kinematic data based on rating joints on a 5-point scale. Using the LSD method on the observed degree value of joint positions, the mean clinicians' gait observations agreed an average of 80% with kinematic gait data.

Our results (mean agreement, 58%) are similar to those found for the EVGS (mean agreement, 64%). Both results were based on percentage agreement of kinematic and estimated category scores of sagittal gait. It is difficult, however, to compare our validity results with the results of other tools when different statistical approaches, such as the Cohen κ, were used. Assessing validity by comparing the agreement between assigned categories can lead to overly optimistic results. This can be the case when an OGA tool uses categories that correspond to a specific range of angular positions and κ scores are based on the agreement of categories, not on agreement of the observed degrees of joint position. There is, therefore, no precise information on how close the observations were to the kinematic data, other than that they lay within the same category range. These ranges of angular values vary between OGA tools, from 15° (OGS) to 20° (VGAS) for most categories, and open-ended categories (ie, >25°) are common to all tools. This means that because of the tools' validity measurements, an observer can deviate by up to 20° (or more in the case of open-ended categories) from the target kinematic data and the validity agreement is still considered to be “perfect.” By using the LSD method for measuring the SF-GT’s validity, we compared the actual observed degrees of joint position with the degrees given by the kinematic gait data. This method makes possible a precise evaluation of the level of validity of the SF-GT and the diversion of the observation from kinematic data.

**Table 2: Level of Agreement, Computed With Percentage Agreements of Categories, Between Mean Observer Data and Kinematic Data for Hip, Knee, and Ankle, and for All Gait Clusters**

<table>
<thead>
<tr>
<th>Gait Cluster</th>
<th>Hip Agreement (mean %)</th>
<th>Knee Agreement (mean %)</th>
<th>Ankle Agreement (mean %)</th>
<th>Mean Agreement (%) for Cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: mild crouch</td>
<td>83*</td>
<td>63</td>
<td>53</td>
<td>66</td>
</tr>
<tr>
<td>2: mobile crouch</td>
<td>62</td>
<td>71</td>
<td>73</td>
<td>69</td>
</tr>
<tr>
<td>3: moderate crouch</td>
<td>39</td>
<td>24</td>
<td>22†</td>
<td>28†</td>
</tr>
<tr>
<td>4: severe crouch</td>
<td>26</td>
<td>82</td>
<td>49</td>
<td>52</td>
</tr>
<tr>
<td>5: mild equinus</td>
<td>80</td>
<td>75</td>
<td>44</td>
<td>66</td>
</tr>
<tr>
<td>6: moderate equinus/knee ext</td>
<td>30</td>
<td>49</td>
<td>24</td>
<td>34</td>
</tr>
<tr>
<td>7: moderate equinus/knee flex</td>
<td>49</td>
<td>74</td>
<td>63</td>
<td>62</td>
</tr>
<tr>
<td>8: severe equinus</td>
<td>78</td>
<td>46</td>
<td>63</td>
<td>62</td>
</tr>
<tr>
<td>9: stiff leg</td>
<td>32</td>
<td>66</td>
<td>65</td>
<td>54</td>
</tr>
<tr>
<td>10: weak plantarflexion</td>
<td>68</td>
<td>80</td>
<td>77</td>
<td>75*</td>
</tr>
<tr>
<td>11: ankle double bump</td>
<td>72</td>
<td>68</td>
<td>44</td>
<td>61</td>
</tr>
<tr>
<td>12: near normal</td>
<td>77</td>
<td>40</td>
<td>64</td>
<td>60</td>
</tr>
<tr>
<td>13: normal</td>
<td>63</td>
<td>52</td>
<td>81</td>
<td>65</td>
</tr>
<tr>
<td>Mean agreement for joint (%)</td>
<td>58</td>
<td>61</td>
<td>56</td>
<td>58</td>
</tr>
</tbody>
</table>

*Highest agreement; †lowest agreement.

**Table 3: Level of Agreement, Computed Using LSD and Degree Estimates, Between Mean Observer Data and Kinematic Data for Hip, Knee, and Ankle, and for All Gait Clusters**

<table>
<thead>
<tr>
<th>Gait Cluster</th>
<th>Hip Agreement (mean %)</th>
<th>Knee Agreement (mean %)</th>
<th>Ankle Agreement (mean %)</th>
<th>Mean Agreement (%) for Cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: mild crouch</td>
<td>73</td>
<td>74</td>
<td>78</td>
<td>74</td>
</tr>
<tr>
<td>2: mobile crouch</td>
<td>68†</td>
<td>69</td>
<td>78</td>
<td>72†</td>
</tr>
<tr>
<td>3: moderate crouch</td>
<td>77</td>
<td>83</td>
<td>78</td>
<td>79</td>
</tr>
<tr>
<td>4: severe crouch</td>
<td>91*</td>
<td>72</td>
<td>84</td>
<td>82</td>
</tr>
<tr>
<td>5: mild equinus</td>
<td>85</td>
<td>77</td>
<td>79</td>
<td>80</td>
</tr>
<tr>
<td>6: moderate equinus/knee ext</td>
<td>83</td>
<td>83</td>
<td>86</td>
<td>84*</td>
</tr>
<tr>
<td>7: moderate equinus/knee flex</td>
<td>77</td>
<td>79</td>
<td>80</td>
<td>79</td>
</tr>
<tr>
<td>8: severe equines</td>
<td>78</td>
<td>83</td>
<td>77</td>
<td>79</td>
</tr>
<tr>
<td>9: stiff leg</td>
<td>76</td>
<td>83</td>
<td>86</td>
<td>82</td>
</tr>
<tr>
<td>10: weak plantarflexion</td>
<td>85</td>
<td>82</td>
<td>82</td>
<td>83</td>
</tr>
<tr>
<td>11: ankle double bump</td>
<td>82</td>
<td>87</td>
<td>81</td>
<td>83</td>
</tr>
<tr>
<td>12: near normal</td>
<td>76</td>
<td>83</td>
<td>81</td>
<td>80</td>
</tr>
<tr>
<td>13: normal</td>
<td>83</td>
<td>83</td>
<td>87</td>
<td>84*</td>
</tr>
<tr>
<td>Mean agreement for joint (%)</td>
<td>80</td>
<td>80</td>
<td>81</td>
<td>Total average 80</td>
</tr>
</tbody>
</table>

*Highest agreement; †lowest agreement.
Assuming that perfect agreement when using observational gait assessment is not achievable, the question of how much deviation from the criterion standard measurement is acceptable remains unanswered. The answer is subjective and is likely to differ, depending on the context for the use of the OGA. The LSD method we used showed statistical significance, which relates only to our confidence that the differences we observed were real and not by chance. Some LSDs were very large (ie, 40°) and certainly beyond any reasonable clinical limits of validity, but they were rare. Large LSDs also occurred when only 1 observer diverted hugely from the kinematic data; this phenomenon reflects the wide-ranging abilities of practitioners to perceive 3-dimensional motion via 2-dimensional images. The majority of joint position observations were assumed that with a more detailed scrutiny is far more detailed than most other OGA tools. The SF-GT includes 18 individual assessments per gait cycle and tools are more commonly assessed by expert clinicians. The further work we described in the development of the SF-GT and its adjustment to facilitate identification of 13 different gait styles defined by prior statistical analysis of kinematic gait data from children with CP and normal gait. Development of existing gait assessment tools is not described in the literature and the use of quantitative kinematic data in tool development is a more systematic approach than what appears to have been adopted in the past. Assessment of the tool’s use by clinicians revealed good agreement between SF-GT scores and quantitative data; consequently, the quantitative and the observation gait data could not have been recorded at the same time. Furthermore, the whole basis of clinical gait assessment is that observations of a small number of cycles can be used to represent a subject’s gait pattern and pathology as it is outside the laboratory. As such, pragmatically it should not matter that we compared data that were not from the same gait cycle, because they are all supposed to represent the same subject and his/her gait pattern. To create a more “like-for-like” comparison, the clinicians’ assessments could have been made using several gait cycles, and this would have taken greater account of the known variations between cycles. This would have proved very time consuming, however, given that the observers took several hours to assess the 13 gait cycles. Also, our use of only 1 cycle means that agreement between SF-GT scores and quantitative data might in fact be better than described here, because the most likely outcome of using only 1 gait cycle is that the agreement is adversely affected.

CONCLUSIONS

We have described the development of the SF-GT and its adjustment to facilitate identification of 13 different gait styles defined by prior statistical analysis of kinematic gait data from children with CP and normal gait. Development of existing gait assessment tools is not described in the literature and the use of quantitative kinematic data in tool development is a more systematic approach than what appears to have been adopted in the past. Assessment of the tool’s use by clinicians revealed good agreement between SF-GT scores and quantitative kinematic data, although the question of what is acceptable disagreement between observations and quantitative kinematic data remains unanswered. Our further work will include the evaluation of the SF-GT’s inter- and intra-rater repeatability.

---

Table 4: Results From 1 Observer Showing the Differences Between Kinematic Data and Observer Data, SD, t Value, Individual LSD, and Mean LSD (all observers) for All Gait Phases of Cluster 5 (mild equinus)

<table>
<thead>
<tr>
<th>Gait Cluster 5: Mild Equinus</th>
<th>Kinematic Data</th>
<th>Observer Data</th>
<th>SD</th>
<th>t=39n</th>
<th>LSD</th>
<th>Mean LSD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hip</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial contact</td>
<td>19</td>
<td>20</td>
<td>0.71</td>
<td>2.02</td>
<td>1.43</td>
<td>31.85</td>
</tr>
<tr>
<td>End double support</td>
<td>18</td>
<td>30</td>
<td>8.49</td>
<td>2.02</td>
<td>17.14</td>
<td>17.14</td>
</tr>
<tr>
<td>Mid stance</td>
<td>−2</td>
<td>10</td>
<td>8.49</td>
<td>2.02</td>
<td>17.14</td>
<td>17.14</td>
</tr>
<tr>
<td>Start double support</td>
<td>−19</td>
<td>−10</td>
<td>6.36</td>
<td>2.02</td>
<td>12.86</td>
<td>13.57</td>
</tr>
<tr>
<td>Toe-off</td>
<td>−3</td>
<td>0</td>
<td>2.12</td>
<td>2.02</td>
<td>4.29</td>
<td>11.28</td>
</tr>
<tr>
<td>Mid swing</td>
<td>27</td>
<td>45</td>
<td>12.73</td>
<td>2.02</td>
<td>25.71</td>
<td>22.21</td>
</tr>
<tr>
<td>Knee</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial contact</td>
<td>20</td>
<td>30</td>
<td>7.07</td>
<td>2.02</td>
<td>14.28</td>
<td>16.07</td>
</tr>
<tr>
<td>End double support</td>
<td>22</td>
<td>20</td>
<td>1.41</td>
<td>2.02</td>
<td>2.86</td>
<td>14.21</td>
</tr>
<tr>
<td>Mid stance</td>
<td>12</td>
<td>10</td>
<td>1.41</td>
<td>2.02</td>
<td>2.86</td>
<td>9.57</td>
</tr>
<tr>
<td>Start double support</td>
<td>14</td>
<td>15</td>
<td>0.71</td>
<td>2.02</td>
<td>1.43</td>
<td>11.28</td>
</tr>
<tr>
<td>Toe-off</td>
<td>61</td>
<td>50</td>
<td>7.78</td>
<td>2.02</td>
<td>15.71</td>
<td>13.43</td>
</tr>
<tr>
<td>Mid swing</td>
<td>71</td>
<td>90</td>
<td>13.44</td>
<td>2.02</td>
<td>27.14</td>
<td>22.07</td>
</tr>
<tr>
<td>Ankle</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial contact</td>
<td>−3</td>
<td>−5</td>
<td>1.41</td>
<td>2.02</td>
<td>2.86</td>
<td>16.00</td>
</tr>
<tr>
<td>End double support</td>
<td>−1</td>
<td>−10</td>
<td>6.36</td>
<td>2.02</td>
<td>12.86</td>
<td>5.50</td>
</tr>
<tr>
<td>Mid stance</td>
<td>1</td>
<td>−10</td>
<td>7.78</td>
<td>2.02</td>
<td>15.71</td>
<td>8.28</td>
</tr>
<tr>
<td>Start double support</td>
<td>2</td>
<td>−5</td>
<td>4.95</td>
<td>2.02</td>
<td>9.57</td>
<td>9.57</td>
</tr>
<tr>
<td>Toe-off</td>
<td>−33</td>
<td>−30</td>
<td>2.12</td>
<td>2.02</td>
<td>4.29</td>
<td>17.85</td>
</tr>
<tr>
<td>Mid swing</td>
<td>−15</td>
<td>−10</td>
<td>3.54</td>
<td>2.02</td>
<td>7.14</td>
<td>15.35</td>
</tr>
</tbody>
</table>
## APPENDIX 1: SALFORD GAIT TOOL—SAGITTAL (SIDE) PLANE VIEW

Enter the observed degrees of ranges of movement in the spaces below. Then assign a CATEGORY from the list on the left for each joint.

<table>
<thead>
<tr>
<th>Name:</th>
<th>Date:</th>
<th>Diagnosis:</th>
<th>Initial Contact</th>
<th>End Double Support</th>
<th>Mid Stance</th>
<th>Start Double Support</th>
<th>Toe Off</th>
<th>Mid Swing</th>
<th>Sum of Category Scores</th>
</tr>
</thead>
</table>

### TRUNK
Circle observation: normal

<table>
<thead>
<tr>
<th>Category</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>backwards</td>
<td>forwards</td>
</tr>
<tr>
<td>backwards</td>
<td>forwards</td>
</tr>
<tr>
<td>backwards</td>
<td>forwards</td>
</tr>
<tr>
<td>backwards</td>
<td>forwards</td>
</tr>
<tr>
<td>backwards</td>
<td>forwards</td>
</tr>
<tr>
<td>backwards</td>
<td>forwards</td>
</tr>
<tr>
<td>Overall:</td>
<td>forwards</td>
</tr>
</tbody>
</table>

### HIP (1)*

<table>
<thead>
<tr>
<th>Category</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>degrees category</td>
<td>degrees category</td>
</tr>
<tr>
<td>2=−21° or more extension</td>
<td>1=−6° to −20° extension</td>
</tr>
<tr>
<td>0=−5° ext to 15° flexion</td>
<td>1=16° to 45° flexion</td>
</tr>
<tr>
<td>2=46° or more flexion</td>
<td></td>
</tr>
</tbody>
</table>

### KNEE (5)*

<table>
<thead>
<tr>
<th>Category</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>degrees category</td>
<td>degrees category</td>
</tr>
<tr>
<td>2=−16° or more extension</td>
<td>1=−6° to −15° extension</td>
</tr>
<tr>
<td>0=−5° ext to 10° flexion</td>
<td>1=11° to 45° flexion</td>
</tr>
<tr>
<td>2=46° or more flexion</td>
<td></td>
</tr>
</tbody>
</table>

### ANKLE (0)*

<table>
<thead>
<tr>
<th>Category</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>degrees category</td>
<td>degrees category</td>
</tr>
<tr>
<td>2=21° or more DF</td>
<td>1=1° to 20° DF</td>
</tr>
<tr>
<td>0=neural 0° to −15° PF</td>
<td>1=−16° to −45° PF</td>
</tr>
<tr>
<td>2=46° or more PF</td>
<td></td>
</tr>
<tr>
<td>Ankle toe strike</td>
<td>Heel off the floor?</td>
</tr>
<tr>
<td>flat foot</td>
<td>Yes</td>
</tr>
<tr>
<td>heel strike</td>
<td>No</td>
</tr>
</tbody>
</table>

Initial Contact = first frame when the foot makes contact with the floor.
End Double Support (loading response) = first frame when toes of opposite foot come off the floor right leg.
Mid Stance = opposite foot is going behind stance leg, toes are just peeping out at the front of stance leg.
Start Double Support (terminal stance) = first frame when opposite foot touches the ground left leg.
Toe Off (initial swing) = first frame when toes (or foot) of front leg have just left the floor.
Mid Swing = first frame when heel of stance leg becomes visible behind swinging heel.

(*) The numbers in brackets show the SUM of CATEGORY SCORES observed for NORMAL gait.

Abbreviations: DF, dorsiflexion; PF, plantarflexion.
### Appendix 1 (cont'd):

**Labels derived from sum of category scores (for sagittal plane)**

<table>
<thead>
<tr>
<th>Joint</th>
<th>Description</th>
<th>Score Range</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hip</strong></td>
<td>Moderate/severe extension</td>
<td>−3 or less</td>
</tr>
<tr>
<td></td>
<td>Mild extension/stiff hip</td>
<td>−2 to 0</td>
</tr>
<tr>
<td>Normal</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Mild/mobile flexion</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Moderate flexion</td>
<td>3 to 6</td>
</tr>
<tr>
<td></td>
<td>Severe flexion</td>
<td>7 or more</td>
</tr>
<tr>
<td><strong>Knee</strong></td>
<td>Severe hyperextension</td>
<td>−1 or less</td>
</tr>
<tr>
<td></td>
<td>Moderate hyperextension/stiff knee</td>
<td>0 to 2</td>
</tr>
<tr>
<td>Normal</td>
<td></td>
<td>3 to 4</td>
</tr>
<tr>
<td></td>
<td>Mild/mobile flexion</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Moderate flexion</td>
<td>6 to 8</td>
</tr>
<tr>
<td></td>
<td>Severe flexion</td>
<td>9 to 10</td>
</tr>
<tr>
<td><strong>Ankle</strong></td>
<td>Severe toe walking/foot drop</td>
<td>−6 or less</td>
</tr>
<tr>
<td></td>
<td>Moderate toe walking/foot drop</td>
<td>−3 to −5</td>
</tr>
<tr>
<td>Normal</td>
<td></td>
<td>−1 to −2</td>
</tr>
<tr>
<td></td>
<td>Mild/mobile dorsiflexion</td>
<td>1 to 2</td>
</tr>
<tr>
<td></td>
<td>Moderate dorsiflexion</td>
<td>3 to 5</td>
</tr>
<tr>
<td></td>
<td>Severe dorsiflexion</td>
<td>6 or more</td>
</tr>
</tbody>
</table>

#### References


**Supplier**
a. Microsoft Corp, One Microsoft Wy, Redmond, WA 98052.

Objective: To evaluate the inter- and intraobserver repeatability of the Salford Gait Tool (SF-GT), a new observational-based gait assessment tool for evaluating sagittal plane cerebral palsy (CP) gait.

Design: Masked comparative evaluation.

Setting: University in the United Kingdom.

Participants: A convenience sample of 23 pediatric physical therapists with varying degrees of clinical experience recruited from the Greater Manchester area.

Intervention: Participants viewed videotapes of the sagittal plane gait of 13 children and used the SF-GT to analyze their 13 different gait styles on 2 occasions. Eleven children had hemiplegic, diplegic, or quadriplegic CP and 2 were neurologically intact.

Main Outcome Measures: Inter- and intraobserver repeatability of hip, knee, and ankle joint positions at 6 different phases of the gait cycle.

Results: The SF-GT demonstrated good interobserver (77%) and intraobserver (75%) repeatability.

Conclusions: We have established that the SF-GT is a repeatable clinical assessment tool with which to guide the diagnosis, treatment planning, and evaluation of interventions by pediatric physical therapists of sagittal plane gait deviations in CP.

Key Words: Cerebral palsy; Gait; Kinematics; Observer variation; Physical therapy techniques; Rehabilitation.

© 2007 by the American Congress of Rehabilitation Medicine and the American Academy of Physical Medicine and Rehabilitation

CLINICAL- AND OBSERVATION-BASED gait examination is fraught with potential inaccuracies. Judgments made from visual observation and subsequent interpretation are subjective and rely heavily on clinicians’ training and experience, which varies widely. If there is a high variation between clinicians in their assessments of gait, then the care pathway a patient enters may depend more on the clinician making the assessment rather than on the true gait problems the patients present with. Several studies have considered whether observation-based gait assessment tools can be used reliably between clinicians or by the same clinician on repeated occasions over time. The Hugh Williamson Gait Laboratory Scale, a modified version of the Physician Rating Scale (PRS), had a low interrater repeatability among 4 experienced raters who rated 25 children with spastic diplegic gait (κ = .46 for foot-strike), although Corry had previously demonstrated an interrater repeatability of κ equal to .67 for the scale’s foot-strike section. Another variant of the PRS, the Observational Gait Scale, had modest interrater repeatability (κ = .58; range, .29–.86) and intrarater repeatability (κ = .69; range, .30–.91) for its first 6 sections. A third variant of the PRS, the Visual Gait Assessment Scale, also had modest interrater repeatability (κ = .67; range, .44–.89) and intrarater repeatability (κ = .53; range, −.04 to .86) for 2 experienced observers. The interrater repeatability for 17 items on the Edinburgh Visual Gait Score (EVGS) ranged from 96% for initial contact to 55% for knee extension in terminal swing (mean interrater repeatability, 70%). Intrarater repeatability was reported to be good for 5 experts in gait analysis, demonstrated by a mean least significant difference of 3.20 (range, 2.63–4.01).

Our previous work established that clinicians have a need for a gait assessment tool for use in routine clinical settings, and we subsequently developed the Salford Gait Tool (SF-GT) with which to assess the gait of children with cerebral palsy (CP). Development of the tool has been described elsewhere and so only an overview of it is presented here. With the SF-GT, users can assess the position of the hip, knee, and ankle at 6 specific events during gait (initial contact, end double support, mid stance, start double support, toe-off, mid swing). Users watch a video of subjects walking and estimate the angular position (in degrees) of each joint at the instant of each of the 6 gait events. The angular position of each joint (recorded in degrees by visual estimation of a video recording) corresponds to a category (2, 1, 0, −1, −2) for that joint at the instant of each of the 6 gait events. There are therefore 6 categories allocated to each joint over the gait cycle and 18 categories in all to describe each lower limb in gait. The sum of the 6 categories for each joint represents the function of the joint over the entire gait cycle. The output from using the tool is therefore a numeric indicator for each of the 3 joints and when the scores for the joints are summed they provide an indication of the entire gait pathology. We have described in related work 13 different gait styles in children with CP (diplegia, hemiplegia, quadriplegia, monoplegia, dystonia, ataxia), based on the statistical analysis of quantitative kinematic data of 56 children with CP and 11 children with normal gait (table 1). The quantitative data used to derive these 13 gait styles were also used to define the boundaries between the 5 categories (2, 1, 0, −1, −2) at the hip, knee, and ankle. This was done so that users of the SF-GT would have the best possible chance of differentiating between the 13 gait styles. Once the tool was developed we sought to assess its intra- and interclinician repeatability prior to its clinical implementation.
METHODS

Participants

Pediatric physical therapists used the SF-GT to assess through video footage the gait of children with CP and normal gait on 2 separate occasions. Twenty-three therapists participated on the first assessment day, of which 17 returned for the second day. Four therapists did not return because of time commitments, 1 therapist completed only 1 set of assessments spread across both days, and 1 set of assessments was spoiled. The observers were recruited from 9 National Health Service Trusts from the Greater Manchester area in the United Kingdom and had a range of clinical experience (table 2). All observers gave written consent prior to participation.

Gait Data

Sagittal plane video recordings of the gait of 13 children (11 children with hemiplegic, diplegic, and quadriplegic CP gait; 2 neurologically intact children) were selected from a gait database comprising 67 children (56 with CP gait, 11 with normal gait). Data had been recorded for prior research using a digital camcorder at 25Hz. That research identified homogenous gait styles of the 67 children based on statistical analysis of quantitative sagittal plane hip, knee, and ankle data. Thirteen different gait styles were identified and the 13 children chosen (mean age, 9y 6mo; range, 6–16y) represented 1 example of each of the 13 different gait styles. This ensured that the observers assessed a wide range of gait styles. Parents of the children had previously given consent for use of the video images for research purposes and the research was approved by both university and health service ethics committees.

Procedure

The observers used the SF-GT to assess the gait of the 13 children on 2 separate occasions 14 days apart. At the first assessment day, they were given a 2-hour update on gait maturation, normal gait kinematics, the characteristics of the different phases of gait, and methods for recording and assessing video recordings of gait. The rationale and history of the SF-GT was also explained, after which there was a demonstration of how to use the SF-GT to assess normal and CP gait (using 2 children with normal gait and 1 child with CP gait). Observers were each allocated a workstation with a DVD player and a television screen. They practiced using the SF-GT together for 30 minutes, during which time they familiarized themselves with its layout and the mechanics of pausing the recording at the appropriate events during gait; the observers also discussed their experiences, which stimulated discussion about the observed joint positions and their interpretations of what they had learned in the update lecture. They informally “calibrated” their observations and interpretations against those of others.

The observers then worked individually without discussion using the SF-GT to assess 1 gait cycle of 1 leg of each of the 13 children. They were permitted to work at their own speed, to review the gait cycles as often as required, and were under no time limit. The assessments were completed in 3 to 5 hours. At assessment day 2, 14 days later, no further training or advice was given. The 17 observers who returned the second day assessed the same video recordings, which were presented to them in an order different from the first day (but, as at the first assessment, the order was the same for all observers). Although there was again no time limit, all assessments were completed within 2 hours 30 minutes.

Data Analysis

Percentage agreements were chosen to describe “exact” agreement between 2 sets of data (assessment 1, assessment 2) and to allow for direct comparison with other studies. The Cohen κ for intraobserver repeatability could not be computed because κ statistics require a symmetric 2-way table in which the first observation uses the same rating categories of the second observation. This was not always the case because some observers’ range of category scores (from ±2 to −2) allocated at the first assessment did not match the range of

<table>
<thead>
<tr>
<th>Physical Therapy Grade</th>
<th>Assessment Day 1</th>
<th>Assessment Day 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Superintendent (most senior)</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Senior I</td>
<td>18</td>
<td>13</td>
</tr>
<tr>
<td>Senior II</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Junior (least senior)</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>23</td>
<td>17</td>
</tr>
</tbody>
</table>
category scores at the second assessment (e.g., assessment 1 range of categories was 1–2, but assessment 2 range of categories was 0–2).

Interobserver repeatability for the assessments was evaluated by calculating the mean percentage agreement at each of the 6 gait events (Table 3). For instance, if 95% of observers scored the hip with category 1 at initial contact, 55% with category 1 at end double support, 100% with category 0 at mid stance, 75% with category 1 at start double support, 100% with category 0 at toe off and 95% with category 1 at mid swing, the mean agreement for this hip would be 87%.

Intraobserver repeatability for each of the 3 joints was evaluated by calculating the percentage agreement at the 6 gait events (Table 4) for each observer. If 6 of the possible 6 category scores for a joint agreed at both assessments, then 100% of the categories agreed. If 5 of 6 scores for a joint were the same, then 83% of the categories agreed. If 4 of 6 scores were the same, then 67% of the categories agreed, and so on for 33% agreement, 17% agreement, and 0% agreement. The intraobserver repeatability for the exact number of degrees allocated to each joint at the 6 phases was also evaluated by counting on how many occasions the exact degree values (e.g., 25° of hip flexion at initial contact) agreed at both assessments for each observer (see Table 4). Intraobserver repeatability for each of the 13 gait styles was evaluated using the mean percentage agreement for the 3 joints from all observers (Table 5). The statistical analysis was performed with SPSS.*

Table 3: Mean Interobserver Percentage Agreement of Joint Categories for the 13 Gait Styles

<table>
<thead>
<tr>
<th>Gait Style (1–13)</th>
<th>Hip Agreement of Category Scores (mean %)</th>
<th>Knee Agreement of Category Scores (mean %)</th>
<th>Ankle Agreement of Category Scores (mean %)</th>
<th>Mean Interobserver Agreement of Category Scores (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: mild crouch</td>
<td>87</td>
<td>85</td>
<td>75</td>
<td>82</td>
</tr>
<tr>
<td>2: mobile crouch</td>
<td>63</td>
<td>89</td>
<td>79</td>
<td>77</td>
</tr>
<tr>
<td>3: moderate crouch</td>
<td>70</td>
<td>80</td>
<td>86</td>
<td>78</td>
</tr>
<tr>
<td>4: severe crouch</td>
<td>64</td>
<td>82</td>
<td>56†</td>
<td>67†</td>
</tr>
<tr>
<td>5: mild equinus</td>
<td>81</td>
<td>85</td>
<td>73</td>
<td>80</td>
</tr>
<tr>
<td>6: moderate equinus/knee extension</td>
<td>66</td>
<td>72</td>
<td>79</td>
<td>72</td>
</tr>
<tr>
<td>7: moderate equinus/knee flexion</td>
<td>82</td>
<td>83</td>
<td>76</td>
<td>80</td>
</tr>
<tr>
<td>8: severe equinus</td>
<td>80</td>
<td>83</td>
<td>67</td>
<td>77</td>
</tr>
<tr>
<td>9: stiff leg</td>
<td>67</td>
<td>82</td>
<td>89</td>
<td>79</td>
</tr>
<tr>
<td>10: weak plantarflexion</td>
<td>83</td>
<td>83</td>
<td>66</td>
<td>77</td>
</tr>
<tr>
<td>11: ankle double bump</td>
<td>81</td>
<td>69</td>
<td>68</td>
<td>73</td>
</tr>
<tr>
<td>12: near normal</td>
<td>91*</td>
<td>77</td>
<td>78</td>
<td>82</td>
</tr>
<tr>
<td>13: normal</td>
<td>90</td>
<td>80</td>
<td>80</td>
<td>83†</td>
</tr>
<tr>
<td>Mean agreement for joint</td>
<td>77</td>
<td>81</td>
<td>75</td>
<td>Total mean 77</td>
</tr>
</tbody>
</table>

*Highest interobserver agreement.
†Lowest interobserver agreement.

Table 4: Mean Intraobserver Percentage Agreement of Joint Categories and Joint Position as Recorded in Degrees for 13 Gait Styles

<table>
<thead>
<tr>
<th>Observer</th>
<th>Hip Agreement of Category Scores (mean %)</th>
<th>Knee Agreement of Category Scores (mean %)</th>
<th>Ankle Agreement of Category Scores (mean %)</th>
<th>Mean Intraobserver Agreement of Category Scores (%)</th>
<th>Mean Intraobserver Agreement of Exact Degrees (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>73</td>
<td>79</td>
<td>81</td>
<td>78</td>
<td>30</td>
</tr>
<tr>
<td>2</td>
<td>79</td>
<td>73</td>
<td>74</td>
<td>76</td>
<td>33</td>
</tr>
<tr>
<td>3</td>
<td>78</td>
<td>77</td>
<td>73</td>
<td>76</td>
<td>41</td>
</tr>
<tr>
<td>4</td>
<td>74</td>
<td>83</td>
<td>74</td>
<td>77</td>
<td>25</td>
</tr>
<tr>
<td>5</td>
<td>63</td>
<td>72</td>
<td>72</td>
<td>69</td>
<td>26</td>
</tr>
<tr>
<td>6</td>
<td>74</td>
<td>76</td>
<td>61</td>
<td>70</td>
<td>ND</td>
</tr>
<tr>
<td>7</td>
<td>72</td>
<td>82</td>
<td>79</td>
<td>78</td>
<td>ND</td>
</tr>
<tr>
<td>8</td>
<td>61</td>
<td>85</td>
<td>82</td>
<td>76</td>
<td>30</td>
</tr>
<tr>
<td>9</td>
<td>67</td>
<td>79</td>
<td>51†</td>
<td>66†</td>
<td>30</td>
</tr>
<tr>
<td>10</td>
<td>78</td>
<td>86</td>
<td>78</td>
<td>81</td>
<td>32</td>
</tr>
<tr>
<td>11</td>
<td>81</td>
<td>83</td>
<td>86</td>
<td>83</td>
<td>33</td>
</tr>
<tr>
<td>12</td>
<td>81</td>
<td>73</td>
<td>69</td>
<td>74</td>
<td>25</td>
</tr>
<tr>
<td>13</td>
<td>82</td>
<td>88</td>
<td>90*</td>
<td>87*</td>
<td>47*</td>
</tr>
<tr>
<td>14</td>
<td>63</td>
<td>69</td>
<td>67</td>
<td>66†</td>
<td>29</td>
</tr>
<tr>
<td>15</td>
<td>77</td>
<td>81</td>
<td>59</td>
<td>72</td>
<td>25</td>
</tr>
<tr>
<td>16</td>
<td>84</td>
<td>75</td>
<td>72</td>
<td>70</td>
<td>26</td>
</tr>
<tr>
<td>17</td>
<td>62</td>
<td>62</td>
<td>81</td>
<td>68</td>
<td>18†</td>
</tr>
<tr>
<td>Mean agreement for observer</td>
<td>72</td>
<td>78</td>
<td>73</td>
<td>Total mean 75</td>
<td>Total mean 30</td>
</tr>
</tbody>
</table>

Abbreviation: ND, no data.
*Highest intraobserver agreement.
†Lowest intraobserver agreement.
Table 5: Mean Intraobserver Agreement (by gait style) of Joint Categories for the 13 Gait Styles

<table>
<thead>
<tr>
<th>Gait Style</th>
<th>Mean Intraobserver Agreement (%) for Gait Style</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: mild crouch</td>
<td>75</td>
</tr>
<tr>
<td>2: mobile crouch</td>
<td>72</td>
</tr>
<tr>
<td>3: moderate crouch</td>
<td>73</td>
</tr>
<tr>
<td>4: severe crouch</td>
<td>62*</td>
</tr>
<tr>
<td>5: mild equinus</td>
<td>77</td>
</tr>
<tr>
<td>6: moderate equinus/knee extension</td>
<td>69</td>
</tr>
<tr>
<td>7: moderate equinus/knee flexion</td>
<td>76</td>
</tr>
<tr>
<td>8: severe equinus</td>
<td>78</td>
</tr>
<tr>
<td>9: stiff leg</td>
<td>74</td>
</tr>
<tr>
<td>10: weak plantarflexion</td>
<td>79</td>
</tr>
<tr>
<td>11: ankle double bump</td>
<td>75</td>
</tr>
<tr>
<td>12: near normal</td>
<td>77</td>
</tr>
<tr>
<td>13: normal</td>
<td>87*</td>
</tr>
</tbody>
</table>

*Highest intraobserver agreement.
†Lowest intraobserver agreement.

RESULTS

Interobserver Repeatability

Between observers, an average of 77% (range, 67%–83%) of hip, knee, and ankle category scores at the 6 phases of gait agreed (see table 3). The observers agreed on 3875 (77%) of a possible 5004 category scores. Of the 1129 scores that disagreed, 98% (1111) differed by 1 category and 2% (18) differed by 2 categories from the mode category score. The knee joint was assessed with greatest repeatability across all observers (mean agreement, 81%), followed by the hip (77%) and the ankle (75%). Across all assessments the highest interobserver agreement was achieved for the hip in gait style 12 (near normal gait) (91%) and the lowest was for the ankle in gait style 4 (severe crouch gait) (56%). Gait style 13 (normal gait) was assessed with the highest overall interobserver repeatability (mean agreement of hip, knee, and ankle, 83%), followed by styles 1 (mild crouch) (82%) and 12 (near normal gait) (82%). Gait style 4 (severe crouch) was assessed with the lowest interobserver repeatability, with 67% mean agreement (see table 3).

Intraobserver Repeatability

Within observers, an average of 75% (range, 66%–87%) of category scores agreed assessed between assessments 1 and 2 (see table 4). Of a possible 3728 category scores, 2792 were identical in both assessments. Of the 936 scores that did not agree between days, 97% (911) differed by only 1 category, 2.4% (22) differed by 2 categories, and 0.3% (3) differed by 3 categories from the mode category score. Within observers, scores for the hip agreed 72%, for the knee 78%, and for the ankle 73%. Exact agreement in the estimated position of the hip, knee, and ankle (recorded in degrees) occurred on average on 30% of occasions (see table 4). This means that almost one third of all degrees estimated at the 2 assessments were identical. Observer 13, however, accounted for nearly half of all exact matches between both assessments (47% of exact matches), and there was marked variation between observers, with observer 17 demonstrating only 18% agreement between assessments. Observers 6 and 7 did not provide degree data (see table 4). Style 13 (normal gait) was assessed with the highest intraobserver repeatability (mean agreement, 87%) while style 4 (severe crouch) was assessed with the least intraobserver repeatability, with 62% mean agreement (see table 5).

DISCUSSION

We found inter- and intraobserver repeatability to be good, and better than with other observational gait assessment tools.1,6–10 The interobserver repeatability was modest for the PRS (κ range, .46–.67) and good for the EVGS (agreement, 70%). The SF-GT’s level of repeatability (interobserver agreement, 77%; intraobserver agreement, 75%) was higher. The SF-GT intraobserver repeatability of the exact position (in degrees) of the joints was better than expected, with nearly one third of all joint positions exactly the same in both assessments. This aspect of assessment repeatability has not been reported previously. The results presented here are particularly pleasing because previous studies involved fewer observers, typically between 2 and 5, and the gait assessment tools used have fewer scoring categories and offer a smaller number of choices or categorical scoring options in terms of describing joint or gait pathology. The larger number of possible scores for each joint and the segmenting of gait into 6 different events means that the SF-GT offers a far greater number of choices than do other tools. We assume this leads to an increased potential for variation between observers and between repeated assessments by the same observer.

The knee joint was assessed with the highest level of agreement between and within observers (between observers, 81%; within observers, 78%). This result agrees with other research4,5,10 that also found the knee joint to be the most repeatable joint to assess visually in the sagittal plane. Observers’ estimations may be facilitated by the clear visibility of the large knee joint, while the femur and the tibia formed 2 natural “goniometric levers.” In contrast, the hip is a less visible joint because soft tissue often obscures the exact pelvis position. Some observers appeared to use the curvature of the spine as a secondary indicator for the pelvis, but this is likely to be a poor indicator of pelvis position. The ankle joint displays the least range of motion between dorsiflexion and plantarflexion (≈30°) and therefore, when coupled with a smaller goniometric lever of the foot, may require more precision during visual estimation.

Normal and near-normal gait styles (normal, mild crouch, near normal) were assessed with higher level of agreement than severe gait abnormalities (severe crouch). This demonstrates that observers had knowledge of and were able to assess normal gait, a prerequisite for assessing abnormal gait. This contrasts with the findings of Eastlack et al5 that their observers were unfamiliar with normative values of gait.

When severe gait abnormalities were present, however, observers demonstrated the least repeatability, possibly because of the complexity of abnormal gait that manifests itself in 3 planes. The majority of the percentage agreements below 70% occurred in hip and the ankle of abnormal gait styles (see tables 3–5). Observations between 2 successive assessments varied between observers, indicating that intraobserver variability is itself variable.

We identified several factors that can affect observer repeatability during the experiment. We used a range of different DVD players, television screens, and liquid crystal display equipment to play back and display the images. We thought this would reflect the reality of using the tool in a clinical setting. Different brands of DVD players, however, appeared to offer more frames than others and their sensitivity to manual operations was different. Observers sometimes found it difficult to find the precise frame they required. We used both flat screen and curved screen television sets and found that joint angles were more difficult to assess on a curved screen. The
precise image frame at which a specific gait event (eg, initial contact) occurred was interpreted differently between observers. There were various levels of fatigue, stamina, and patience between observers. They made variable allowances for errors in the 2-dimensional image of the leg, particularly in cases where there was medial (internal) rotation at the hip that distorted the users’ perspective of sagittal plane joint motion. Some observers attempted to adjust their estimation of joint angles according to the perceived amount of rotation. They made clinical judgments based on the gait problems they identified and those judgments may have affected their objectivity in estimating joint angles. For example, when a child was toe walking the clinical judgment would be that the Achilles’ tendon was likely to be shortened and the ankle would consequently be plantar flexed, while in fact the “toe walking” was caused by increased hip and knee flexion with the ankle being near its 90° (neutral) position. Errors (including errors in basic arithmetic) were made while in the process of assigning grades to the joints. None of these factors are specific to the SF-GT, but have not been mentioned previously.

Our results are also pleasing because they were obtained in a scenario that might not be conducive to attaining highest possible levels of repeatability. The research design was pragmatic and sought to reflect some realities of undertaking gait assessment in a clinical setting. Participants were not experts in gait assessment and none were routinely involved in quantitative gait analysis. As noted, there were several practical issues raised during the experiment about the use of equipment and interpretation of images. Experience in the training sessions suggested that knowledge of normal gait values and terminology was limited among some observers. Other studies have typically evaluated assessment tools using “experts”6-8-10 which we assume leads to higher intra- and interrepeatability.

Observer repeatability could be improved by using markers on bony landmarks to indicate joint centers, by using flat-screen televisions or personal computers. Training with peer support would seem essential in ensuring appropriate implementation of an observation-based clinical gait assessment tool. Using goniometers to measure joint angles on the display screen was suggested as a future possibility by our observers to enhance validity and repeatability of observations. Considering the problems associated with poor repeatability of using goniometers,15,16 however, a strict protocol would be required. Also, users may begin to rely too heavily on the goniometers when, for example, errors in the recording process, or medial hip rotation, distort the 2-dimensional image on the screen. Subjectivity would still remain. Alternatively, technology could be developed that would automatically extract segment angles from video images and remove the subjective “visual” element from gait assessment.

CONCLUSIONS

The SF-GT demonstrated good inter- and intrauser repeatability, comparable and better in some cases than reports of other clinically orientated observation-based gait assessment tools. The evaluation presented here was generally more pragmatic than evaluations of other tools, involving more a detailed analysis, more observers, and the use of nonexperts to test the tool. Despite good results, problems associated with the subjective nature of observational gait assessment and variations between clinicians remain. The use of the SF-GT to assess sagittal plane motion in CP gait would be complementary to an examination of gait in the frontal and transversal planes, as well as a full clinical examination of joint mobility, muscle power, and tone. There are some practical issues that could be addressed, but training and experience are key aspects that have not been fully explored. In future work we will evaluate whether the SF-GT has the sensitivity to detect treatment effects and changes in gait over time.

References


Supplier

a. Version 11.5; SPSS Inc, 233 S Wacker Dr, 11th Fl, Chicago, IL 60606.
Estimated Prevalence of Obstructive Sleep Apnea–Hypopnea Syndrome After Cervical Cord Injury

Bernard E. Leduc, MD, FRCP, Jehan H. Dagher, MD, FRCP, Pierre Mayer, MD, FRCP, François Bellemare, PhD, Yves Lepage, PhD


Objectives: To estimate the prevalence of obstructive sleep apnea–hypopnea syndrome (OSAHS) in patients with cervical cord injury and to identify predictive factors.

Design: Cross-sectional study.

Setting: Rehabilitation center.

Participants: Forty-one adults with cervical cord injury of more than 6 months in duration.

Interventions: Medical history, physical exam, and full in home overnight polysomnography were undertaken. Data were collected on characteristics of spinal cord injury, current medication, sleeping habits, daytime sleepiness, body mass index (BMI), and neck circumference.

Main Outcome Measure: Presence or absence of OSAHS as defined by the American Academy of Sleep Medicine criteria (1999).

Results: Twenty-two (53%) patients (95% confidence interval [CI], 38.4%–68.9%) had OSAHS. Daytime sleepiness (odds ratio [OR], 41.1; 95% CI, 2.3–739.7; P<.02), BMI of 30 kg/m² or higher (OR=17.2; 95% CI, 1.4–206.4; P<.03), and 3 or more awakenings during sleep (OR=34; 95% CI, 1.6–744.8; P=0.03) were the best predictive factors of OSAHS obtained by a forward stepwise multiple logistic regression.

Conclusions: The estimated prevalence of OSAHS is high after cervical cord injury. OSAHS should be suspected, especially in patients with daytime sleepiness, obesity, and frequent awakenings during sleep.

Key Words: Rehabilitation; Sleep apnea syndromes; Spinal cord injuries; Tetraplegia.

© 2007 by the American Congress of Rehabilitation Medicine and the American Academy of Physical Medicine and Rehabilitation

Sleep disorders are frequently encountered in persons with a spinal cord injury (SCI), who also suffer from various problems such as neurogenic pain, spasticity, and other medical conditions related to post-traumatic spinal cord lesions. Sleep apnea, mainly the obstructive type (and more rarely central or mixed), is among the sleep disorders associated with cervical cord injury, resulting in repeated interruptions of breathing during sleep which, when accompanied by daytime symptoms, defines the presence of obstructive sleep apnea–hypopnea syndrome (OSAHS). The pathogenesis of OSAHS in patients with tetraplegia remains obscure, but potential risk factors have been described, such as neck thickness, change in the normal ventilatory response to hypoxic stimuli, possible thickening of the oropharyngeal wall through unopposed parasympathetic stimulation of the mucosa and vessels wall (blocking of sympathetic spinal input), longer sleep periods spent in the supine position than that observed in the general population, a tendency to obesity, and the use of baclofen.

Sleep studies without regard to symptoms have established the prevalence of sleep apnea and hypopnea in middle-aged adults to be 24% in men and 9% in women. According to a recent investigation in subjects with tetraplegia, however, the prevalence reaches 62% (95% confidence interval [CI], 32%–86%). The prevalence of OSAHS in the general population, taking into accountsymptomatic subjects, is 4% in men and 2% in women whereas in patients with SCI, the rates are found to be higher, reaching 9% and up to 48%. The large variety of methodologies used among studies no doubt accounts for the disparity between published rates and makes comparisons difficult. For instance, among some compiled studies, diagnostic criteria were provided by various parameters (arterial oxygen saturation level, airflow, pulse, respiratory efforts) without night-time polysomnography. Others investigated a limited sample size ranging from 10 patients to 16 patients. The selection of participants also varied with regard to age of the subjects selected (restricted to more than 40 years old in 1 study or the inclusion of some but not all groups of SCI subjects classified according to the criteria of the American Spinal Injury Association (ASIA).

Because of wide variations in the reported occurrence of OSAHS in persons with tetraplegia, we thought it would be beneficial to estimate more precisely the prevalence of OSAHS after cervical cord injury using the diagnostic criteria for OSAHS recommended in 1999 by the American Association of Sleep Medicine (AASM). To our knowledge, this study has never been done before in this population. If a high estimated prevalence of OSAHS in cervical SCI is confirmed by these criteria, clinicians could be made more aware of this fact, and could be better prepared to detect it and consequently treat it, in these very patients for whom quality of life and life expectancy are already less than optimal.

The main objective of this exploratory study was to estimate the prevalence of OSAHS in subjects with a post-traumatic tetraplegia according to the diagnostic criteria recommended by the AASM; a secondary goal was to identify predictive clinical factors of OSAHS in this population.
Participants

After approval of this research project by the Research Ethics Committee, we recruited participants at the Outpatient Clinic of the Institut de Réadaptation de Montréal without any prior advertising or postings and without selecting patients on the basis of symptoms that could suggest sleep apnea; our intention was thus to avoid a possible bias in assessing the prevalence of OSAHS. Patients were invited to participate in the study during their medical examination if they fulfilled the following criteria: age 18 years or more with a complete or incomplete cervical cord injury for more than 6 months. Excluded were patients with a history of sleep apnea, chronic cardiorespiratory disease, or requiring mechanical ventilation. Because the sleep technician had to drive to set up the equipment for in-home sleep studies and return the following morning to pick up the equipment, those living more than 25km outside the city of Montreal, QC, and who did not agree to commute to the hospital for 1 night of polysomnography, were excluded. The sample size goal was a minimum of 40 participants, similar to the few studies with this high number of participants. Age, SCI duration, level of the SCI (C4 to C8), and the degree of impairment according to the ASIA Impairment Scale (grades A, B, C, or D) were documented.

Medical History

Medical history focused on the symptoms that are required for the diagnosis of OSAHS according to the AASM: excessive daytime sleepiness, daytime fatigue, impaired concentration, choking or gasping during sleep, unrefreshing sleep, and recurrent awakenings from sleep. If present, the degree of sleepiness was assessed according to the Epworth Sleepiness Scale (ESS). The items were summed with a maximum score of 24. An average score of 5.9±2.2 is obtained in the general healthy population and a score of more than 10 is considered abnormal. The number of self-reported awakenings from sleep was also recorded.

Although not included in the AASM diagnostic criteria for OSAHS, presence or absence of snoring (occasionally or regularly) and the use of baclofen or benzodiazepines were recorded, because these factors have been in some studies associated with OSAHS after SCI.

Clinical Examination

We recorded the following physical examination elements: blood pressure, height, and weight to compute the body mass index (BMI). Healthy weight is defined by a BMI ranging from 20 to 24.9kg/m², overweight by a BMI of 25 to 29.9kg/m², and obesity by a BMI 30kg/m² or more. We also included neck size, which is found to correlate very well with the BMI and the frequency of sleep apnea.

Polysomnography

Thirty-nine participants underwent standard polysomnography during 1 night of sleep in their home (unsupervised), and 2 subjects were recorded at the hospital with the same equipment.

The study included an electroencephalogram (leads at: C3 and A2, C4 and A1, O1 and A2, A2 and O2), a submental electromyogram and an electromyogram of the anterior tibialis muscles, a bilateral electro-oculogram, a measure of oronasal airflow (pressure transducer), oxygen saturation by pulse oximeter (SpO₂), an electrocardiogram, body position recording (sides, supine, prone) in percentages during sleep, and a measure of respiratory efforts using thoraco-abdominal straps (strain gauges). Total sleep time was measured as well as sleep efficiency, expressed as the percentage of total sleep time minus periods of awakening (electroencephalogram) during sleep. A minimum recording period of 4 hours was required for a valid sleep study in accordance with standards for sleep studies.

Diagnostic Criteria for OSAHS

We based the diagnosis of OSAHS on polysomnography criteria associated with either excessive daytime sleepiness (criterion A) or 2 or more of the following that are not better explained by other factors (criterion B): recurrent awakenings, choking or gasping during sleep, unrefreshing sleep, daytime fatigue, and impaired concentration. Excessive daytime sleepiness was defined according to the International Classification of Sleep Disorders as difficulty in staying awake and involuntarily falling asleep in the daytime, which could not be better explained by other factors. Overnight polysomnography shows 5 or more obstructed breathing events per hour during sleep lasting 10 seconds or longer (obstructive apnea or hypopnea event) or respiratory effort-related arousals of 10 or more seconds duration per hour of sleep. Apnea was defined as a complete oronasal airflow interruption; hypopnea corresponded to a decrease of at least 50% of this flow or a reduction of less than 50% if associated with greater than 3% oxygen desaturation or arousal. Arousals were scored according to the American Sleep Disorders Association criteria. They were classified as spontaneous or associated with respiratory efforts, snoring, or periodic leg movements.

The number of apneas and/or hypopneas per hour of sleep defined the apnea-hypopnea index (AHI). Depending on the frequency of events, the OSAHS was classified as mild (AHI score, 5–14), moderate (AHI score, 15–30), or severe (AHI score, >30).

Sleep was scored by experienced polysomnographic technicians according to Rechtshaffen and Kales criteria and all tracing reviewed by the sleep physician.

Statistical Analysis

We report data on continuous variables as mean ± standard deviation (SD) with minimum and maximum values (range), whereas data on discrete variables are reported according to relative frequency. The prevalence rate was assessed by a 95% CI. The groups with or without OSAHS were compared using a Pearson chi-square test for discrete variables and a Student t test (with Satterthwaite-Welch correction for heterogeneous variances) for continuous variables. To further assess motor complete higher level cervical SCI as a factor associated with OSAHS, participants with a C4 or C5 SCI (ASIA grade A or B) were compared with participants with a motor incomplete lower-level lesion C6, C7, or C8 (ASIA grade C or D). After dichotomizing 19 clinical variables, the 8 clinical variables for which the P value of the Pearson chi-square test was lower than .25 were used in a stepwise forward logistic regression to determine the ones that best predict OSAHS. The Pearson correlation coefficient was also used to measure the linear relationship between continuous variables. Significance levels were set at 5% (P<.05) for all analyses. The software used was Stata.

RESULTS

Participants

From the 73 subjects approached, 15 declined to participate, and 4 were excluded for various reasons (1 subject already diagnosed with OSAHS, 2 lived too far, 1 had a multiresistant bacterial infection). Of the 54 subjects recruited, 13 did not complete the study, in 1 case because of too few total sleep time (<4 h), and 12 subjects abandoned the study (loss of interest or additional illness translating into refusal to submit to polysomnography). Therefore, 41 participants were studied; etiology of trauma was motor vehicular in 14 (34%) cases, work in 8 (19%), sports and recreation in 13 (32%), and falls in 6 (15%). Main characteristics of the 41 participants are reported in Table 1.

Comparison of the sociodemographic and medical characteristics (sex, age, SCI level and duration) between this group and the group of the 13 subjects who did not complete the study showed no significant differences.

Prevalence of OSAHS

Flow signal was reliable for all polysomnographic recordings, because no segments showed unusable data. Among the 41 participants, 5 or more episodes of apnea/hypopnea per hour of sleep were observed in 23 (56%) subjects, of whom 22 (17/22 men, 5/8 women) also fulfilled the diagnostic criteria for OSAHS: the estimated prevalence of OSAHS was thus established to be 53% (95% CI, 38.4–68.9%).

Criteria A and B were both present in 16 of the 22 apneic participants, whereas sleepiness (criterion A) was seen in 2 subjects, and 4 others were both present in 17 subjects. With the same approach without taking into account the number of awakenings during sleep, that is, looking only at sleepiness and the BMI, the rate of predictability made it possible to correctly reclassify 78.5% of subjects. With the same approach but without taking into account BMI or apneas, the rate of predictability was 70% of the subjects.

Sleep Study Results

AHI reached 27.5±28.3 (range, 5–96) in the OSAHS group and 5±4.3 (5–19) in the non-OSAHS group. In the OSAHS group, the average duration of apneas and hypopneas lasted, respectively, 26±9.8 (11–40) seconds and 29.5±11.5 (14–52) seconds. Arousals per hour of sleep, more frequent in the OSAHS group (13±8.8 vs 7.4±4.9), were mostly related to respiratory efforts or to snoring episodes. Sleep efficiency and its overall structure did not differ significantly between both groups.

Factors Associated With OSAHS

The following clinical factors were associated with OSAHS: BMI (P = .02), obesity (P = .02), and neck size (P = .04). Comparison of the groups on the basis of symptoms showed that they differed only by excessive daytime sleepiness (Table 2). Sleepiness reported by apneic subjects did not seem to be related to the intake of medications (P = .12).

The average ESS score in OSAHS subjects (7.9±5.5) was significantly greater (P = .03) than in the group without OSAHS (4.8±3.4), but an ESS score greater than 10 was unable to significantly differentiate the 2 groups of participants (P = .1). Finally, we did not find a correlation between ESS score and the AHI value (r = .23, P = .28).

No association was established between OSAHS and the other variables measured: age, sex, SCI level, group according to the ASIA Impairment Scale, SCI duration, snoring, fatigue, benzodiazepines, or baclofen intake average daily dose of 53 mg (OSAHS) and 75 mg (no OSAHS).

Predictive Model of OSAHS

Three dichotomous clinical variables were obtained by a forward stepwise multiple logistic regression to predict OSAHS successfully: sleepiness (presence, absence), the BMI <30 kg/m², ≥30 kg/m², and the number of awakenings during sleep (<3, ≥3). Thus, the predictive model, which included these 3 variables, sleepiness (odds ratio [OR], 41.1; 95% CI, 2.3–739.7; P = .02), BMI ≥30 kg/m² (OR = 34; 95% CI, 1.6–744.8; P = .03), and 3 or more awakenings (OR = 17.2; 95% CI, 1.4–206.4; P = .03), made it possible to correctly reclassify 78.5% of subjects. With the same approach but without taking into account the number of awakenings during sleep, that is, looking only at sleepiness and the BMI, the rate of predictability was 70% of the subjects.

DISCUSSION

Considering our study population representative of the general population of our area, because all patients from this population with an SCI from all causes are treated at the Institut
de Réadaptation de Montréal, the current study suggests a high prevalence of OSAHS in subjects with a cervical SCI compared with the general population and concurs with previous investigations that actually prove comparable in terms of sample type and size and the use of polysomnography.\textsuperscript{7,9,12,15,21} Nevertheless, given the relatively small sample size in all these studies, including ours, the results must be interpreted with caution. This study may not estimate the prevalence of OSAHS in persons with a cervical SCI accurately, but when compared with previous studies, our results, like others, are suggestive of a high prevalence of OSAHS.

Even though 5 of 7 women were diagnosed with OSAHS, this small number of female participants did not allow us to assess if male sex predominance was an associated factor with OSAHS as reported in the general population.

The majority of studies that did not use polysomnography as a criterion for OSAHS reported lower rates, demonstrating the greater sensitivity of polysomnography in confirming an OSAHS diagnosis. The results obtained from the 39 unsupervised home polysomnography are considered valid because it has been shown that unattended full polysomnography can be performed in the home with reliable and high-quality recordings, relative to the values in hospital settings.\textsuperscript{30}

The reasons underlying the prevalence of OSAHS in subjects with tetraplegia are still unknown, but given the list of potential factors briefly listed in the introduction, it is likely that the causes are multifactorial.

Obesity is a factor associated with OSAHS in the general population, but its rate (22%) in our sample was not greater than in the population at large. This obesity, observed in 9 participants, 8 of whom had OSAHS, is most likely enhanced by a decrease in daily energy expenditure and in basal metabolism.\textsuperscript{8} Furthermore, given the decrease in lean muscle mass (muscle atrophy) and the increase in adipose tissue, it is possible that the BMI probably underestimates the real percentage of fat in subjects with tetraplegia\textsuperscript{31} and therefore the actual number of obese subjects within this population.

The possibility that some of the participants already suffered from OSAHS prior to their SCI cannot be totally disregarded despite a review of their personal and familial history. Though the Multivariate Apnea Prediction Index\textsuperscript{32} questionnaire is of value in identifying the presence of OSAHS prior to trauma, it was not used because the length of the post-traumatic time period prior to the study was thought to be too long (average, 14y), which could lead to a loss of reliability in answering the questionnaire.

Though sleeping in the supine position in able-bodied persons is sometimes linked to OSAHS and does increase the AHI, our study did not show any correlation between this variable and OSAHS because the majority of participants in both groups slept mainly on their back.

As also reported by other authors,\textsuperscript{7,12,14,16} we found no association between OSAHS and baclofen, a spasmolytic agent similar to the inhibitory neurotransmitter, \(\gamma\)-aminobutyric acid, which is known for its decreasing effects on ventilation and response to hypercapnea and hypoxemia.\textsuperscript{33} But contradictory conclusions have also been published.\textsuperscript{3} Most probably, a careful comparison of the results published with regard to the treatment of subjects with a cervical SCI with baclofen, that is, the daily doses used, length of treatment and level of SCI, would yield more precise conclusions.

The possibility of a selection bias must be contemplated and could have contributed to the high prevalence. Even though the study had not been publicized prior to the patient recruitment, it remains possible that some subjects decided to participate for personal reasons or in response to complaints about sleep disorders from their partners, resulting in some recruitment bias. Despite this last hypothesis, it is still true that if we were to consider that the 22 subjects with a diagnosis of OSAHS were the only ones out of the 72 approached (73 minus the apneic patient) at the outpatient clinic, the prevalence of OSAHS would then be 30%, and this value would still be way beyond that found in the general population. A case control study would prevent this selection bias insofar as the solicitation of subjects and their basic demographic characteristics would be comparable for both groups. Such a study, but retrospective (from looking at medical records), has already emphasized some factors associated with OSAHS in persons with a cervical SCI,\textsuperscript{34} and a prospective control study would seem advisable to bypass this difficulty and shed more light on the issue.

Among the symptoms looked for as diagnostic criteria for OSAHS, excessive daytime sleepiness, observed in 19 of 22 apneic subjects, is the most frequent. This symptom is nonspecific, however, because 10 nonapneic patients were also somnolent and sleepiness has been described in 30% to 50% of the general population without OSAHS.\textsuperscript{35} The presence of sleepiness seems to be more useful for the diagnosis of OSAHS than the definition of excessive sleepiness provided by a score greater than 10 on the ESS,\textsuperscript{23} the threshold “diagnostic” number described in the literature in an able-bodied population. In our study, a score of more than 10 failed to distinguish between the groups with or without OSAHS despite a significant difference between their average scores. No validation evidence for the ESS has been reported in the SCI population, however. Furthermore, both the practicability and the reliability of the ESS score seem questionable because patients frequently underestimate the importance of their somnolence\textsuperscript{32} and correlation between the ESS score and the AHI is weak.\textsuperscript{36}

The precision level of the described predictive model for OSAHS leads us to recommend its use for persons with tetraplegia while remaining cautious given the nonspecificity of the clinical factors underpinning it. It therefore seems advisable that clinicians treating patients with a cervical SCI be sufficiently informed about OSAHS to introduce in their practice some elements that could allow them to diagnose OSAHS when present. OSAHS would then be less likely to be underdiagnosed, and the appropriate treatment\textsuperscript{37} applied more quickly.

**Study Limitations**

This study has some limitations. Apart from the already mentioned possibility of a selection bias, we must remember that a risk factor not looked for in our study, the presence of cranio-orofacial anomalies, could have been underestimated.

Another limitation is the relatively small sample, comparable with other similar studies, which reduces the statistical power of our results. Validation of the proposed predictive model is needed to assess the clinical relevance of the findings in patients with tetraplegia. Finally, this study makes use of the 1999 AASM task force diagnostic criteria for OSAHS and changes to the OSAHS definition have since been published.\textsuperscript{38}

**CONCLUSIONS**

Using specific diagnostic criteria, the results of our study in 41 participants with post-traumatic tetraplegia suggest a high prevalence of OSAHS, much greater than in the general population. OSAHS should be suspected and looked for, particularly in patients with tetraplegia presenting daytime sleepiness, obesity, and frequent awakenings during sleep.
Acknowledgments: We thank Vincent Jobin, MD, FRCP, Respiratory Division, Centre Hospitalier de l’Université de Montréal (CHUM), for his contribution to the study analysis, Benoit Dulong, Department of Mathematics and Statistics, Université de Montréal, for the data analysis, and the 3 sleep technicians who carried out the polysomnographic recordings: Marie-France Gilbert, Fanny Morin, and Manon Robert. The editorial assistance of Ovid Da Silva, Research Support Office, CHUM Research Centre, is acknowledged.

References


Suppliers

a. Suzanne portable recording system; Tyco Healthcare, 15 Hampshire St, Mansfield, MA 02048.
b. Version 8; StataCorp, 4905 Lakeview Dr, College Station, TX 77845.
Low-Frequency Rectangular Pulse Is Superior to Middle Frequency Alternating Current Stimulation in Cycling of People With Spinal Cord Injury

Johann Szecsi, MD, MSc (Eng), Ché Fornusek, PhD, Phillip Krause, MD, Andreas Straube, MD


Objective: To determine the efficacy of using modulated middle frequency alternating current (MFAC) muscle stimulation for functional electric stimulation–propelled cycling by people with spinal cord injury (SCI) compared with the conventional method of using standard low-frequency rectangular pulses (LFRP).

Design: Repeated-measures.

Setting: Laboratory setting.

Participants: Eleven otherwise healthy volunteer subjects with SCI (8 with American Spinal Injury Association [ASIA] grade A, 3 with ASIA grade B).

Interventions: To evaluate cycling-relevant differences between LFRP and modulated MFAC stimulation, we exposed participants to isometric measurements and cycling experiments performed during both 20Hz LFRP and 4kHz modulated with 50Hz MFAC.

Main Outcome Measures: We recorded maximal isometric torque, maximal dynamic work during 20 minutes of ergometer cycling, and perceived discomfort for each of the 2 stimulation patterns.

Results: Both the isometric torque (P<.02) and work generated (P<.001) during MFAC stimulation were significantly lower than during standard LFRP stimulation. Four participants reported discomfort and 1 of them also developed skin burns during MFAC stimulation.

Conclusions: Our findings suggest that in SCI subjects, stimulated cycling with low frequency is generally more effective than cycling with modulated MFAC in terms of torque, work, and pain sensation.

Key Words: Electric stimulation; Rehabilitation; Spinal cord injuries; Torque; Work.

© 2007 by the American Congress of Rehabilitation Medicine and the American Academy of Physical Medicine and Rehabilitation
highly fatigable fast-twitch fibers (IIa, IIb). Additionally, because SCI subjects have decreased sensation, maximal fiber recruitment is often achievable because muscle stimulation intensity is not limited by pain.

Two recent observations suggested to us that MFAC stimulation could offer improvement over LFRP in muscle torque and power outputs during FES-propelled cycling. First, we learned of the case of an exceptionally well-trained FES cyclist who had recovered from paraplegia after a severe injury. This case was shared with us by a local physiotherapist. The subject had been able-bodied for several years and had a 4-channel stimulator that provided continuous 4kHz sinusoid current. Second, recent research showed that fatigue rate varies with frequency in able-bodied subjects. Because of the selective dropout of fast-twitch, fast-fatigue (high fatigue rate) fibers at higher kilohertz frequencies, there is a greater proportional contribution of slow-twitch, fatigue-resistant (low fatigue rate) motor units at these frequencies, which is the fatigue rate reduces monotonically in the 1 to 10kHz range. Fatigue rate reduction could hypothetically improve functional outcome.

Because maximum electrically induced torque in able-bodied subjects decreases monotonically with increasing frequency (at 15kHz the maximum torque is ~50% of that at 1kHz), the reduced fatigue rate at higher MFAC frequencies is apparently achieved at the cost of reduced maximum evocable force. Studies with able-bodied subjects have suggested that this trade-off would only be important for people with severe atrophy, in whom it is necessary to evoke maximum muscle force to induce the greatest strength and hypertrophy gains. In such cases the MFAC frequency could be lowered to increase force. The selective dropout of type IIb muscle fibers (fast-twitch, fast-fatigue) at kilohertz stimulation frequencies occurs because in able-bodied subjects the fatigue rate of type IIb fibers is much greater (orders of magnitude) than type I muscle fibers. In contrast, chronically paralyzed muscle consists primarily of type IIb and type Ia fibers whose comparative fatigue rates only differ by a factor of 50 to 60. It is not known whether a selective dropout of fast-fatigue muscle fibers type Ila takes place in SCI subjects similar to what occurs in the able-bodied.

We questioned whether MFAC stimulation of the weak (typical torque range, 7%–10% of MVC) and rapidly fatiguing leg musculature of SCI subjects could achieve a balance between fatigue rate and force to yield a greater functional outcome of movement than in the classic LFRP stimulation. Therefore, our goal in this research was to compare the effectiveness of MFAC and standard LFRP stimulation in producing FES cycling. The functional outcome parameters of FES cycling, which we deemed important and which we measured, were the evoked maximal short time isometric torque, work generated in a fixed time interval (instead of fatigue rate), and pain sensation.

**METHODS**

**Participants**

Eleven otherwise healthy people with chronic SCI (8 with American Spinal Injury Association [ASIA] grade A, 3 with ASIA grade B) and low levels of muscle spasm (Modified Ashworth Scale score range, 0–2) participated in this study (table 1). The muscle fiber composition of their paralyzed muscles was stable. For this study, subjects who had limited experience in FES cycling training (average of 6–28mo for 0.6 training sessions per week) were selected. Therefore, their strength condition corresponded to the typical initial situation of SCI patients joining the outpatient clinic. The University of Munich ethics committee approved the study and the subjects gave their written informed consent prior to their participation.

**Study Design**

Each subject underwent 3 different experimental sessions: (1) isometric measurements using LFRP and MFAC stimulation, (2) ergometry using LFRP stimulation, and (3) ergometry using MFAC stimulation. Session order was randomized and each session was performed on a different day within 6 weeks. Subject 8 did not participate in the isometric measurement.

**Stimulation**

The quadriceps, hamstrings, and glutei muscle groups were electrically stimulated for ergometer cycling. Pairs of auto-adhesive gel electrodes (Flextrode) were placed on the skin over the proximal and distal fourth of each muscle bulk. A constant current 8-channel stimulator (Motionstim) provided the LFRP current (rectangular, biphasic, charged balanced pulses; frequency, 20Hz; maximum pulse amplitude, 127mA; constant pulse width, 500μs) (fig 1A). For MFAC stimulation a middle frequency constant current 6-channel stimulator provided the modulated MFAC (4kHz sinusoidal modulated with 50Hz on-off rectangles; duty cycle, 1:1) (fig 1B). We selected a carrier frequency of 4kHz because it had previously produced a pronounced reduction of fatigue rate in able-bodied subjects.

The middle frequency stimulator could provide maximally 140V peak-to-peak. Assuming a skin resistance of 1kΩ, at 140V the current density on the electrode area (1.4mA RMS/cm², where RMS is root mean square) did not exceed the safety limit for alternating current (ie, 2mA RMS/cm²). Both stimulators had the capacity to reach the saturation region of the recruitment curves for all participants. During the isometric measurements and the ergometer cycling the stimulators were controlled from a laptop computer by serial communication (fig 2).

During ergometer cycling, the laptop directed the muscle stimulator to induce muscle contractions at the appropriate crank angles (table 2) to produce pedaling. Because muscle stimulation angles were statically defined, angular compensations were necessary for the dynamic application. Assuming a muscle force rise time of 140ms, it was calculated that at 35rpm muscle stimulation should occur 28° earlier.

**Isometric Torque Measurements**

A stationary tricycle with its front wheel replaced by a torque transducer served as the test bed for isometric torque measurements (see fig 2). An 8-bit incremental encoder, syn-
chronized to turn with the crankshaft, determined the actual position of the crank. The ankle joint was immobilized at 90° and leg movement was restricted to the sagittal plane by using shank and foot orthoses. Shanks and feet were fixed to the orthoses by self-adhesive (Velcro) straps. A lever manually moved the crank into 18 equiangular positions. First, the maximal torque-producing crank position was determined for each muscle group; the optimal crank angle (e.g., for LFRP stimulation of the left quadriceps of subject 10 the optimal crank angle was 75°, with corresponding left hip angle of 58° and knee angle of 93°) (see fig 2), maximal torque, and maximal current were noted. Stimulation at 40% to 80% of the maximal current (depending on the patient) was then sequentially applied to the 6 muscle groups. The stimulation phases lasted 1.5 seconds and were followed by breaks of 3 seconds duration.

Moments were evoked at the given crank angle by consecutively using LFRP and MFAC stimulation (electrode leads were connected alternately to the low- and the middle-frequency stimulators by means of a switch) and were subsequently recorded on the laptop. Torque recordings of the 6 muscle groups provided the isometric moment versus angle characteristics.

We attempted to include only active muscle contributions to the torque recordings by eliminating offline the bias of isometric torque due to passive gravitational and elastic components. Furthermore, the isometric moments were extrapolated to maximal torque (100%) values taken at optimal crank angles. Conforming to common FES cycling practice, only the positive (acting in the drive direction) half-waves of the moment versus angle characteristics were used to calculate the sum of these half-waves (see Results and figs 3A, 3B). The sum was integrated over 0° to 360° and divided by 360° to give the mean maximal isometric torque, which was used in further processing. We used the mean maximal isometric torque in addition to maximal quadriceps torque because in FES cycling the mean maximal torque, including quadriceps, hamstrings, and glutei positive half-wave contributions, must overcome the drive resistance.

Ergometric Experiments

Because untrained SCI subjects are quite weak, we used an ergometer (see fig 2) with a motor-powered brake and drive. This permitted a low minimal braking torque of 1Nm on the crank (in the first gear). We measured the tangential forces applied by the rider’s right and left legs using strain-gauge instrumented cranks that could measure up to 1kN, and had an accuracy of ±1N. The participant sat on the ergometer chair in the same geometrical position (hip–crank axis distance and tilt) as when isometric measurements were taken on the cycle.

All participants had to complete 20 minutes (considered in a previous study to be relevant for SCI cycling) of continuous pedaling. After a warm-up phase of several minutes, during which the legs were passively turned by the ergometer motor, the stimulation intensity was gradually increased over an average 5 ± 1.7 minutes to the maximum current (100%). Ergometer cycling resistance was increased with stimulation intensity to maintain pedaling cadence in the range of 35 to 55rpm. Stimulation was fixed at 100% for the remainder of the pedaling. As a subject’s muscles fatigued and cadence dropped below 36rpm the braking resistance was manually reduced (in steps of 0.7Nm) to restore pedaling cadence. The crank angle position (10-bit resolution) and the tangential forces collected by the instrumented crank arm were transmitted serially every 50ms to the laptop for recording.

Cadence was computed from the change of crank position over time, which was digitally filtered with a second-order Butterworth filter with a cutoff frequency of 4Hz. The resultant 2-legged crank torque was calculated by adding the left and right tangential crank forces, multiplying this sum by the crank arm length (1.5m), and smoothing it by a second-order, zero-phase low-pass filter with a time constant of 1 second. Pedaling work was computed as the integral of cadence by crank torque over the 20 minutes of stimulation (20-min work, see Results and figs 4A, 4B), which included the stimulation startup period (increasing stimulation) and maximal stimulation period. We did not consider the warm-up period of passive cycling for integration.

During ergometer cycling at maximal stimulation the visual analog scale (VAS) was recorded to subjectively quantify the discomfort of each stimulation pattern.

Statistical Analysis

The sample size of 11 was chosen according to a required minimal detectable difference of 3600J (estimated in prelim-
Fig 2. Isometric and ergometer measurement setups. Because patients' geometric sitting position and the stimulator setup were identical, a combined draft of isometric and ergometer experiments is presented. The crank angle was set by manually turning the torque transducer axle (TTA) by a lever and fixing it with a screw during isometric measurement. Resistance torque and crank angle provided by TTA and decoder (C), respectively, were collected. Electrode leads were connected at a given crank angle alternately to the LFRP and MFAC stimulator by a switch. During ergometric measurements pedaling was motor braked. Tangential forces provided by built-in sensors in the crank arm and crank angle were collected. Either LFRP or MFAC stimulation was used. Legend: 1, 2, 3: stimulation of the quad, hamstrings, and gluteus muscle groups, respectively.

RESULTS

Representative sample data collected during isometric measurements of subject 10 are shown in figures 3A and 3B. The LFRP mean isometric torque showed a clear advantage over MFAC mean isometric torque (14.9Nm vs 13.6Nm). The slight differences in shape of the torque profiles result from the different tissue depths reached by LFRP and MFAC stimulation. Moreover, the comparative analysis of the power courses collected in ergometric measurements of a representative strong subject (no. 10) and a representative weak subject (no. 9) (figs 4A, 4B, respectively) showed that the 20-minute pedaling work during LFRP is also superior to that during MFAC stimulation (13717J vs 7007J and 8152J vs 6001J, respectively). The participants' mean isometric torques, 20-minute pedaling work, and VAS pain scores are presented in table 3. Isometric measurements were not made with MFAC stimulation for subject 8 because she refused further participation after developing blisters during the MFAC ergometer cycling. A comparison of LFRP and MFAC stimulation conditions using descriptive statistics (fig 5) gave maximal isometric torques of 16.6±10.6Nm and 14.2±10.0Nm (n = 10), respectively, and a 20-minute work of 8445±5552J and 4716±1834J (n = 11), respectively. The minimum detectable difference for 20-minute work was 3729J; this was lower than the prework estimated difference, therefore fulfilling the requirement of power greater than 0.7.27

The nonparametric Wilcoxon paired-sample test proved that the isometric torque elicited during standard LFRP stimulation was significantly greater (± [sum of negative ranks] = 8, P < .02), and also that the 20-minute pedaling work generated during LFRP was highly and significantly greater (t = 0, P < .001) than during MFAC stimulation.

Our analysis of the 20-minute work data suggests the hypothesis that the advantage of LFRP over MFAC stimulation intensifies with increasing work (see figs 3A, 3B). To test this hypothesis, we performed a Spearman rank-correlation test of the work difference between the 2 stimulation conditions. We found a significant rank correlation (ρ = .87, P < .002) between

<table>
<thead>
<tr>
<th>Muscle Group</th>
<th>Start Patient 10 (deg)</th>
<th>Stop Patient 10 (deg)</th>
<th>Start All 11 Subjects (deg)</th>
<th>Stop All 11 Subjects (deg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left quadriceps</td>
<td>349</td>
<td>171</td>
<td>7±22</td>
<td>183±11</td>
</tr>
<tr>
<td>Left hamstrings</td>
<td>87</td>
<td>272</td>
<td>94±31</td>
<td>259±41</td>
</tr>
<tr>
<td>Left gluteus</td>
<td>87</td>
<td>272</td>
<td>94±31</td>
<td>259±41</td>
</tr>
<tr>
<td>Right quadriceps</td>
<td>177</td>
<td>0</td>
<td>191±28</td>
<td>23±25</td>
</tr>
<tr>
<td>Right hamstrings</td>
<td>320</td>
<td>85</td>
<td>287±46</td>
<td>83±23</td>
</tr>
<tr>
<td>Right gluteus</td>
<td>320</td>
<td>85</td>
<td>287±46</td>
<td>83±23</td>
</tr>
</tbody>
</table>

NOTE. Values are mean ± standard deviation (SD) unless otherwise indicated. These stimulation angle ranges contain an angular compensation of 28°. Hamstring and gluteus muscle ranges were set equal due to the stimulator's technical limitations. Zero degrees refers to the backward-pointing left crank arm.
the intercondition difference and the absolute amount of generated 20-minute work with LFRP stimulation. In contrast, there was no significant rank correlation between the intercondition difference and absolute amount of elicited isometric moment ($r = 0.05$, $P > 0.5$).

While only subject 4 complained of discomfort during LFRP stimulation, 4 participants reported an abdominal tugging discomfort during MFAC stimulation (located above the neurologic level and in the zone of partial preservation). The Wilcoxon paired-sample test revealed no stimulation mode-conditioned significant differences in the VAS pain sensation during maximally stimulated ergometer cycling ($t = 0$, $P = 0.13$). The descriptive statistics indicated that LFRP appeared to be more comfortable than MFAC stimulation (0.25±0.81 vs 0.9±1.35). Moreover, subject 8 developed blisters at the stimulation sites with MFAC stimulation.

**DISCUSSION**

**Torque and Work**

The major finding of this study is that LFRP stimulation is superior to MFAC in producing isometric torque and dynamic cycling work by paralyzed skeletal leg muscle of SCI subjects. We expected that LFRP isometric torque would be superior to MFAC torque. This finding is similar to findings in studies with able-bodied subjects (we know of no studies on MFAC stimulation-evoked torque or work in SCI subjects).
A second important finding is that pedaling work produced during the first 20 minutes of cycling is higher during maximal stimulation with LFRP than with MFAC. A comparison of LFRP and MFAC conditions showed that 20 minutes of pedaling work diminished more (down to average of 59% [4716 J/8445 J]) than maximal isometric torque (down to an average of 86% [14.2 Nm/16.6 Nm]). This means that the reduction of the maximal torque that can be evoked during MFAC is accompanied by an increased fatigue rate in SCI patients, which is in contrast to MFAC stimulation in able-bodied subjects, where there was a decreased fatigue rate at increasing kilohertz frequencies.

**Explanation of the Increased Dynamic Fatigue Rate in MFAC Stimulation**

Any interpretation of these findings must consider that the rate of fatigue during voluntary dynamic exercise (shortening contractions) in able-bodied persons (no similar data are available for SCI) is greater than during voluntary isometric contraction, that is, the dynamic fatigue rate is generally higher than the static fatigue rate. Presumably the reasons for the greater fatigue rate in dynamic contractions are: (1) the energy requirements are higher for dynamic compared with isometric exercise, and (2) shortening contractions are more sensitive to the metabolic changes that occur in exercising muscle and they alter the force-velocity relationship through the selective fatigue of fast-twitch fibers. This effectively transforms the muscle into a slower type with reduced power output. Moreover, this could be explained by the complete exhaustion of the fast-twitch fibers during dynamic contractions, which has been observed in cat muscles at higher contraction velocities.

Our data suggest that in chronically paralyzed muscle, MFAC stimulation induces more dynamic fatigue (vs isometric fatigue) than LFRP stimulation, thus enhancing the effects of the 2 general causes of dynamic fatigue by using a common physiologic pathway. Two mechanisms may be involved in accentuating dynamic fatigue: (1) high-frequency stimulation, particularly MFAC, is energetically more demanding than LFRP stimulation, and (2) the dropout of fast-fatigue IIb fibers in MFAC stimulation actually signifies a "conversion" into a slow-fatigue IIa type of muscle with a reduced power output. This is similar to the fast-fiber dropout during MFAC stimulation of able-bodied subjects.

**Advantage of LFRP Over MFAC Stimulation is Strength Condition Dependent**

This study has proven that the difference between LFRP and MFAC work correlates highly significantly with the absolute amount of work produced by an individual subject. Therefore, as regards induced functional movement, the advantage of LFRP stimulation over MFAC stimulation becomes more important as an SCI subject’s strength increases. We assume that strength is indicated by the absolute amount of work performed.

From a practical viewpoint, our results show that the force contribution produced by fast-fatigue IIb fibers (or simply by more fibers in case of nonselective dropout) cannot be ignored when optimizing FES cycling in people with SCI. LFRP has the most advantage over MFAC stimulation in strong SCI subjects, who probably have more fast-fatigue IIb fibers (or simply more fibers if nonselective dropout applies) than do weaker subjects. We expect that the well-trained FES cyclist that we mentioned earlier, whose performance was outstanding, could probably drive even better with LFRP than with MFAC stimulation.

**Pain**

We measured pain in SCI subjects during the application of relevant maximal stimulation. In contrast to studies of the able-bodied, there was a nonsignificant trend for more pain and discomfort to be perceived when MFAC stimulation was used. It was remarkable that all participants who still had pain sensations (3 of 4 had sensory incomplete SCI) felt abdominal tugging during MFAC stimulation. Because MFAC stimulation penetrates more deeply into the quadriceps and glutei muscles, it may affect deeper branches of the iliohypogastric nerve. High current densities that cause thermal burns (blisters) are also more likely to occur with MFAC stimulation than with standard LFRP stimulation.

**Table 3: Mean Maximal Isometric Torques, 20-Minute Pedaling Work, and Pain Score (on the VAS) for Study Participants**

<table>
<thead>
<tr>
<th>Patient No.</th>
<th>LFRP Torque (Nm)</th>
<th>MFAC Torque (Nm)</th>
<th>LFRP Work (J)</th>
<th>MFAC Work (J)</th>
<th>LFRP VAS Score</th>
<th>MFAC VAS Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>35.0</td>
<td>28.7</td>
<td>22620</td>
<td>7140</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>2</td>
<td>14.5</td>
<td>11.0</td>
<td>8887</td>
<td>6300</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>3</td>
<td>11.9</td>
<td>6.1</td>
<td>4115</td>
<td>2703</td>
<td>0.0</td>
<td>1.6</td>
</tr>
<tr>
<td>4</td>
<td>6.6</td>
<td>6.5</td>
<td>3213</td>
<td>2503</td>
<td>2.7</td>
<td>3.8</td>
</tr>
<tr>
<td>5</td>
<td>8.4</td>
<td>7.1</td>
<td>7304</td>
<td>5087</td>
<td>0.0</td>
<td>2.2</td>
</tr>
<tr>
<td>6</td>
<td>15.1</td>
<td>15.8</td>
<td>7179</td>
<td>2804</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>7</td>
<td>13.2</td>
<td>7.9</td>
<td>4228</td>
<td>2729</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>8</td>
<td>14.5</td>
<td>*</td>
<td>4919</td>
<td>3897</td>
<td>0.0</td>
<td>2.3</td>
</tr>
<tr>
<td>9</td>
<td>11.6</td>
<td>8.9</td>
<td>8152</td>
<td>6001</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>10</td>
<td>14.9</td>
<td>13.6</td>
<td>13717</td>
<td>7007</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>11</td>
<td>37.3</td>
<td>40.0</td>
<td>8760</td>
<td>5702</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

NOTE. Boldface denotes nonzero VAS values. *Measurement not performed.
Further Work

Comparison of the functional outcome achieved in SCI subjects with 4kHz MFAC and LFRP stimulation raises the question of whether there is a carrier frequency in the range of 1 to 4kHz that would ensure a function outcome superior to that obtained from LFRP. Based on our results and a comparison with isometric torque performed in able-bodied subjects, it can be assumed that the isometric torque in SCI subjects is higher and the dynamic fatigue rate is lower at 2.5kHz MFAC than at 4kHz MFAC. It therefore seems expedient to determine in further investigations whether 2.5kHz MFAC is more effective than LFRP stimulation in regard to torque, work, and dynamic fatigue rate.

The correlation between the interconditional work difference LFRP-MFAC and subject strength that we found in this study suggests that LFRP stimulation-based training of subjects with SCI increases the advantage of LFRP over MFAC stimulation in regard to work performed. Stefanovska and Vodovnik found that MFAC stimulation-based training of the able-bodied led to less isometric force gains than LFRP stimulation-based training. It must still be determined whether MFAC stimulation-based training of SCI subjects can efficiently strengthen the IIa fibers (slow-fatigue) through the presumed selective dropout of the Iib fibers (fast-fatigue). In our opinion, this might explain the outstanding performance described in our SCI case study.17

CONCLUSIONS

Cycling stimulated with LFRP appears to be generally more effective than with 4kHz MFAC in terms of functional outcome (ie, torque, work, and pain sensation) in subjects with complete or incomplete SCI.

Acknowledgment: We thank Judy Benson for copyediting the manuscript.

References


Suppliers
a. Krauth+Timmermann Ltd, Poppenbütteler Bogen 11, D-22399 Hamburg, Germany.

b. Trainer 6 channel stimulator; ETI Ltd, Ingenieurbüro für Medizintechnik, Am Sandfeld 4, D-76149 Karlsruhe, Germany.
c. Funtrike; Noviconsult Ltd, Keferloher Marktstr 23, D-85640 Putzbrunn/Solailinden, Germany.
d. T30FN torque wave; Hottinger Baldwin Messtechnik Ltd, Am Tiefen See 45, D-6100 Darmstadt 1, Germany.
e. Motomed viva 2; Reck-Technik Ltd, Reckstr 1-4, D-88422 Betzenweiler, Germany.
f. Powertec pedal sensor; O-tec Ltd Kraft- und Leistungssysteme, Nibelungenstr 209, D-64625 Bensheim, Germany.
g. Version 6.1.0; The MathWorks Inc, 3 Apple Hill Dr, Natick, MA 01760-2098.
Effects of Osteoarthritis and Fatigue on Proprioception of the Knee Joint

Meral Bayramoglu, MD, Reyhan Toprak, MD, Seyhan Sozay, MD


Objective: To evaluate the impact of knee osteoarthritis (OA) and periarticular muscular fatigue on knee joint kinesthesia.

Design: Cross-sectional study.

Setting: A physical medicine and rehabilitation outpatient clinic.

Participants: Fifty patients with bilateral OA of the knee, and a control group of 30 age-matched healthy volunteers.

Interventions: Not applicable.

Main Outcome Measures: The Kellgren-Lawrence grading system was used to determine the radiographic severity of knee OA. The Lequesne index of severity for knee osteoarthritis was used for assessment of pain, kinesthesia was measured by determining angle reposition error at the knee joint using isokinetic dynamometry, and muscle strength was measured by isokinetic dynamometry.

Results: Reposition errors did not differ between the patient and the control groups, nor did they differ between pre- and postexercise.

Conclusions: Mild-to-moderate OA of the knees does not affect reposition error. Fatigue produced by mild-to-moderate exercise also has no effect on reposition error.

Key Words: Fatigue; Kinesthesis; Osteoarthritis; Proprioception; Rehabilitation.

© 2007 by the American Congress of Rehabilitation Medicine and the American Academy of Physical Medicine and Rehabilitation

ALTHOUGH THE DEFINITION of proprioception is fairly difficult, we have defined it as the ability to detect, without visual input, the spatial position and/or movement of limbs in relation to the rest of the body. It serves to protect against injurious movement and it is critical to the maintenance of joint stability.1 It is also important for normal joint coordination during movement.2 Researchers have tested the sense of both joint position and joint motion, or kinesthesia, to assess proprioception. Joint position sense has been tested by placing subjects’ legs in various predetermined angles of flexion, then asking the subjects to reproduce their perception of the angle of flexion on a visual analog scale.3-4 Joint position sense has also been assessed by having subjects reposition their legs in a remembered angle of flexion.5-9 To test the sense of joint motion, or kinesthesia, researchers have examined the point at which patients can detect slow passive motion. This point is measured in degrees of angular displacement, and is called the threshold to detection of passive motion.5,6,10,11 Both kinesthetic sense and visual analog model have been used in one study to assess proprioception.5 Knee mechanoreceptors promote stability by providing sensory feedback. It has been reported in the literature that age,1,3,6 muscular fatigue,12 and articular disease such as osteoarthritis (OA),7,8 can all have negative effects on proprioception. OA can cause changes that affect not only intracapsular tissues, but also periarticular tissues such as ligaments, capsule, tendons, and muscle, leading to proprioceptive deficits both in extremes of joint position and in body position.13,14 These effects on proprioceptive sense may induce errors in the normal coordinated patterns of the muscles, thereby causing disturbances in functional stability. In this sense, poorer muscle function may contribute to the severity and progression of the disease.15 Muscle function can be evaluated by several means, including manual muscle strength tests, electromyography, magnetic resonance imaging, and isokinetic tests. It is not known, however, whether reduced proprioception causes degenerative arthritis because of reduced muscle reflex, or is actually caused by degenerative arthritis. Studies have shown that kinesthetic sensitivity is decreased by fatigue16,17 and increased by long-term physical practice.18-20 Among these studies, effects of fatigue on shoulders16 and knee18 joint movement sense were assessed, whereas in the other studies,17,19,20 effects of fatigue on position sense of the elbow and knee joints were examined. Animal studies have shown that some proprioceptive receptors are affected by muscle fatigue19-21 and/or by increased intramuscular concentrations of substances released during muscle contractions.22-24 It is assumed that these receptors would be similarly affected in humans, but little is known about how fatigue actually affects human proprioception.

The objective of this study was to further investigate how OA of the knees and periarticular muscular fatigue affect the movement sense, or kinesthesia at the knee joint, which is recognized to be a separate submodality of proprioception. The method we used to measure kinesthesia differed somewhat from previous studies. It was not the reproduction of a remembered angle of joint flexion, nor it was the threshold to detection of passive motion, but it was a combination of the 2 methods.

METHODS

The study subjects were 50 outpatients with bilateral OA of the knees, and 30 healthy volunteers. The dominant extremity was the right extremity in all subjects. OA was diagnosed according to the criteria established by the American College of Rheumatology. The severity of OA in each affected knee was radiographically graded using the Kellgren-Lawrence grading system,22 and the Lequesne index of severity for knee osteoarthritis (ISK)22 was used to quantify pain. Patients with very severe pain who reported that they were not independent in activities of daily living because of the pain in their knees and those with grade 4 OA were excluded because
they were unable to exercise on a cycle ergometer. Other exclusion criteria included knee OA secondary to another disease, any musculoskeletal disease other than OA, any lower-limb joint replacements, or any neurologic conditions (eg, Parkinson’s disease, stroke).

Ethics approval for the study was obtained from the ethics committee of the university and subjects gave informed consent.

We recorded age, sex, height, weight, body mass index (BMI), and quadriceps and hamstring muscle strength (and ratios comparing these 2 muscle groups) for each of the 66 subjects. For the 50 patients, we also recorded the OA grade for each knee and presence and absence of mediolateral instability with valgus and varus stress tests in the right and left knees. In addition, reposition error was tested bilaterally in all controls and in all but 1 patient before and after a fatiguing exercise. This 1 patient had recently undergone arthroscopic surgery on the right knee and underwent reposition error testing of the left knee only.

We used a computerized isokinetic dynamometer⁴ to assess reposition error. For this, the subject was seated on the bench of the dynamometer with hips and knees flexed at 90°. Inflatable pressure boots were worn to immobilize the feet and thus eliminate any sensation cues from the skin or from ankle positioning. Knee range of motion was set at 0° to 90° and the continuous passive-motion mode of the dynamometer extended and flexed the knee at 5°/s. Before starting to take measurements, the physician pushed the “stop” button when the knee was at the middle of the range (target angle for the study, 45°) of the movements from flexion to extension, and from extension to flexion. This was repeated 5 times, and then the patient was instructed to, with eyes closed, push the button when he/she sensed the knee was at 45° of flexion. The reposition angle was recorded 5 times during flexion to extension, and 5 times during extension to flexion. The reposition angles for the first flexion-to-extension movement and first extension-to-flexion movement were omitted, and the mean for the other 4 reposition errors in each movement category was recorded. Reposition error was defined as the difference between the target angle and the reposition angle, and the absolute value of this error was used for statistical analysis. To evaluate the effect of fatigue on reposition error, each subject was asked to push the button when he/she sensed the knee was at 45° of flexion again after moderate exercise (5min on the cycle ergometer). Subjects pedaled at 35 to 45rpm, their maximum heart rate not exceeding 100 beats per minute for the moderate exercise protocol. This protocol was chosen to standardize the exercise of the patients are shown in table 2. The isometric strength of the quadriceps and hamstring muscles was measured using the isokinetic dynamometer with the subject’s hips and knees flexed at 90°. The ratio of quadriceps strength to hamstring strength (Q/H ratio) was also calculated for each subject’s left and right sides.

We used SPSS software⁵ for Windows for statistical analysis. Two-way analysis of variance (ANOVA) was used to compare numerical values within and between groups. Analysis of reposition error in relation to OA grade was made by using the Student t test for independent samples. Relationships between parametric variables were analyzed using the Pearson correlation coefficient. P values less than .05 were considered statistically significant. Intraclass correlation coefficients (ICCs) as described by Fleiss⁶ were used to test the reliability of the proprioception measurements. ICC values above .75 represent excellent reliability; values between .40 and .75 represent fair-to-good reliability; and values below .40 indicate that the testing was unreliable.

RESULTS

The patient group comprised 39 (78%) women and 11 (22%) men. The control group consisted of 25 women (83%) and 5 men (17%). The mean age, mean BMI, and sex distribution of the patient and control groups are shown in table 1. There was no statistically significant difference between the groups with respect to age or BMI.

Seventeen (34%) patients exhibited instability in both knees, 2 (4%) exhibited instability of the left knee only, and 7 (14%) exhibited instability of the right knee only. The mean ISK score ± standard deviation (SD) for the patient group was 7.45 ± 3.16 (range, 1–14). The radiographic OA grades of the patients are shown in table 2.

The results for isometric quadriceps and hamstring muscle strength revealed that both muscle groups were significantly weaker in the patient group than in the control group. The groups’ muscle strength data are shown in table 3.

Testing showed that our process of assessing proprioception using isokinetic dynamometry had fair-to-good reliability. The ICCs for all measurements ranged from 0.5 to 0.7 (table 4).

Table 2: Radiographic Gradings for the Patient Group

<table>
<thead>
<tr>
<th>Grade</th>
<th>Right Knee</th>
<th>Left Knee</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade 0</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Grade 1</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>Grade 2</td>
<td>25</td>
<td>24</td>
</tr>
<tr>
<td>Grade 3</td>
<td>11</td>
<td>13</td>
</tr>
<tr>
<td>Grade 4</td>
<td>ND</td>
<td>ND</td>
</tr>
</tbody>
</table>

Abbreviation: ND, no data (ie, no patients had this grade).

Table 3: Mean Isometric Muscle Strengths and Mean Q/H Ratios for the Patient and Control Groups

<table>
<thead>
<tr>
<th>Strengths and Ratios</th>
<th>Patient Group</th>
<th>Control Group</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>R quadriceps (Nm)</td>
<td>2.44 ± 0.92</td>
<td>3.03 ± 1.18</td>
<td>.015</td>
</tr>
<tr>
<td>L quadriceps (Nm)</td>
<td>2.25 ± 0.86</td>
<td>2.84 ± 1.31</td>
<td>.018</td>
</tr>
<tr>
<td>R hamstring (Nm)</td>
<td>0.63 ± 0.39</td>
<td>0.82 ± 0.44</td>
<td>.055</td>
</tr>
<tr>
<td>L hamstring (Nm)</td>
<td>0.61 ± 0.36</td>
<td>0.74 ± 0.43</td>
<td>.147</td>
</tr>
<tr>
<td>R Q/H</td>
<td>4.89 ± 3.01</td>
<td>4.34 ± 1.78</td>
<td>.371</td>
</tr>
<tr>
<td>L Q/H</td>
<td>4.49 ± 2.33</td>
<td>4.39 ± 2.03</td>
<td>.993</td>
</tr>
</tbody>
</table>

NOTE. Values are mean ± SD. Abbreviations: L, left; R, right.
Statistical analysis using 2-way ANOVA did not reveal any difference between group results and pre- versus postexercise results. The data of the 2 groups for reposition errors before and after exercise are shown in Table 5.

Analysis of reposition error in relation to OA grade revealed no significant differences in reposition error either before or after exercise between the subgroups of patients with grade 1 and grade 2 OA, or between the subgroups with grade 2 and grade 3 OA. The grade 1 subgroup, however, showed significantly smaller reposition error than the grade 3 subgroup at both time points (from flexion to extension, \( P = .008 \); from extension to flexion, \( P = .005 \)) for the right knee. For the left knee, reposition errors did not differ between the grade 1 subgroup and grade 3 subgroups of patients. The reposition errors for the subgroups with grade 1 and grade 3 OA for the right and left knees are shown in Table 6.

Table 4: ICCs of the Pre- and Postexercise Measurements Performed in 20 Subjects

<table>
<thead>
<tr>
<th>Measurement</th>
<th>ICC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Right knee flexion to extension, pre-exercise</td>
<td>.673</td>
</tr>
<tr>
<td>Right knee extension to flexion, pre-exercise</td>
<td>.719</td>
</tr>
<tr>
<td>Left knee flexion to extension, pre-exercise</td>
<td>.573</td>
</tr>
<tr>
<td>Left knee extension to flexion, pre-exercise</td>
<td>.681</td>
</tr>
<tr>
<td>Right knee flexion to extension, postexercise</td>
<td>.499</td>
</tr>
<tr>
<td>Right knee extension to flexion, postexercise</td>
<td>.535</td>
</tr>
<tr>
<td>Left knee flexion to extension, postexercise</td>
<td>.526</td>
</tr>
<tr>
<td>Left knee extension to flexion, postexercise</td>
<td>.693</td>
</tr>
</tbody>
</table>

Table 5: Mean Pre- and Postexercise Reposition Errors* for the 2 Groups and the Tests of Between-Subjects Effects

<table>
<thead>
<tr>
<th>Knee</th>
<th>Group</th>
<th>Pre-Exercise</th>
<th>Postexercise</th>
</tr>
</thead>
<tbody>
<tr>
<td>R flex-ext</td>
<td>Patient</td>
<td>6.29±4.12</td>
<td>6.11±4.81</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>5.20±3.80</td>
<td>5.68±2.94</td>
</tr>
<tr>
<td>L flex-ext</td>
<td>Patient</td>
<td>6.45±4.01</td>
<td>6.90±4.59</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>5.90±2.84</td>
<td>6.47±2.72</td>
</tr>
<tr>
<td>R ext-flex</td>
<td>Patient</td>
<td>8.95±5.94</td>
<td>8.66±3.49</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>7.76±3.91</td>
<td>6.18±3.89</td>
</tr>
<tr>
<td>L ext-flex</td>
<td>Patient</td>
<td>8.11±5.95</td>
<td>7.25±4.82</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>8.55±2.65</td>
<td>6.92±3.30</td>
</tr>
</tbody>
</table>

NOTE. Values are mean degrees ± SD. The results of 2-way ANOVA. The statistical significance of the tests for group (patient vs control), pre- vs postexercise and right vs left knee is underlined in the data below:

Abbreviations: ext, extension; flex, flexion.

*Tests of Between-Subject Effects Dependent Variable: Reposition Error

Corrected model: 683.687\(^{1}\), \( df = 15 \), Mean Square: 45.579, \( F = 2.374 \), Signif: .002, Power: 0.986

DISCUSSION

The results of this study indicate that reposition error in patients with OA of the knees is similar to that in controls with similar age. We did find, however, that patients with more severe OA based on radiographic grading had poorer proprioception than those with less severe radiographic signs. Some of the previous studies that have investigated the relationship between knee OA and proprioception have shown a direct relationship between knee OA and reduced proprioception,\(^{3,7,31,32}\) and some have examined how the combination of age and OA of the knees affects proprioception.\(^{2}\) In line with our findings, a previous study\(^{33}\) documented lower proprioceptive acuity in patients with severe OA than in patients with mild OA.

Some research has shown that proprioceptive deficits are correlated with Western Ontario and McMaster Universities Osteoarthritis Index scores.\(^{1}\) Sharma et al\(^{34}\) found that patients with unilateral knee OA had poorer proprioception than controls in both knees with no clinical or radiographic evidence of OA in either knee. Based on these findings, they speculated that impaired proprioception does not exclusively result from local disease in knee OA.

In most of the above-mentioned previous studies, the severity of knee OA was radiographic grade 2 or higher. In our study, we had to exclude patients with grade 4 OA because their severe pain precluded exercise on the cycle ergometer. In correlation analysis revealed that reposition error was not associated with age, sex, BMI, left or right quadriceps strengths, left or right hamstring strengths, or left or right Q/H ratios in either the patient or control group. As well, in the patient group, neither ISK score nor knee instability correlated with reposition error (\( P > .005 \)).
contrast with most other studies, our investigation included many knees with grade 1 disease. Thus, our study compared the reposition errors of patients with mild-to-moderate bilateral OA of the knees with the reposition errors of normative controls. Had we included patients with more severe OA, we might have observed significantly greater reposition error in the group with OA. In fact, we found that the grade 3 OA patients had greater right-knee reposition error than the patients with grade 1 OA at both time points. The data do not provide any explanation for why this would apply only to the right knee, however. This could speculatively be related to dominant side.

Research has shown that periartricular muscle strengthening exercise has a clear positive effect on proprioception. In our study, we found that the muscles supporting the knee joint were weaker in the patient group than in the control group. Speculatively, reflex inhibition because of pain might be responsible for this weakness, or this group might not be exercising as much as a non-OA subject population, thus leading to weaker muscle strength. Analysis showed, however, that neither muscle strength nor ISK score nor mediolateral knee instability was correlated with reposition error in the OA patients. This can be explained by the presence of factors such as proper muscle and ligament balance, pain, and proprioception, all of which play individual roles in the pathogenesis of OA but also influence each another. Taking all these factors into consideration, reposition error, as we assessed in our study, may not be different between groups, but still there might be proprioceptive deficits in the OA group, because we cannot assume that only reposition error will reflect any deficit in proprioception. As well, none of our patients had grade 4 OA, and mild-to-moderate disease is another likely reason why we found no relationship between muscle strength and reposition error.

There have also been conflicting findings with respect to the effects of periartricular muscular fatigue on proprioception. Marks and Quinney found that proprioception was not affected by a fatigue protocol involving concentric and eccentric quadriceps contractions. In another study where shoulder proprioception was measured in healthy subjects after light and hard exercise, the acuity of movement sense was reported to be significantly reduced in the setting of localized muscle fatigue, and that the effect was more pronounced with hard exercise, and in women. In our study, we observed no reduction of reposition error in either the patient group or control group after fatiguing exercise. Indeed, there was a trend toward smaller reposition error after exercise, but the difference was not statistically significant. This partial but not significant increase in proprioception could have been a result of learning. As well, the exercise in our study was only of moderate intensity, and this could also explain this result. Our OA patients could not tolerate intense exercise for more than 5 minutes. To standardize the conditions in both groups, we had all subjects perform 5 minutes of mild-to-moderate intensity exercise on the cycle ergometer. A limitation of the study here is that we did not use the Borg rating of perceived exertion or any percentage of maximum voluntary contraction of muscle groups. We preferred to produce a standardized fatigue based on time, pedaling rate, and heart rate. In this procedure, there may have been patients and/or subjects who performed the exercise, but yet were not truly fatigued. The ways in which fatigue affects proprioception could be more clearly evaluated in a group of healthy volunteers who can tolerate harder and longer exercise, and by using a more definite way of producing fatigue.

Different studies have assessed proprioceptive acuity using different methods. Threshold of detection of passive movement has been used for assessment of joint position sense in the normal and pathologic knee joint. The method of passive joint position setting and analog reproduction using a goniometer has been used, with reproducible results, in other studies. In our study, we did not test the threshold to detection of passive knee motion, nor did we ask subjects to reproduce a remembered angle of flexion. Our method of measuring proprioception, as described before, was in between the 2 other methods used in previous studies, and the reliability analysis revealed that the reliability was not excellent, but was fair to good.

CONCLUSIONS

Our data revealed no difference in reposition error between patients with bilateral knee OA and healthy controls. The results, however, do suggest that patients with more advanced (grade 3) knee OA have greater reposition errors of the right knees than those with less severe (grade 1) disease. Investigations that include patients with more severe (grade 4) OA might show reduced proprioceptive acuity compared with normative controls. Mild-to-moderate intensity exercise does not affect reposition error in patients with knee OA or in healthy subjects. Indeed, although not significant, there might be a trend toward smaller error (which would indicate better proprioception, not worse) after exercise, because of learning.

References

KINESTHESIA IN OSTEOARTHRITIS AND FATIGUE, Bayramoglu


Suppliers

a. Cybex 770 Norm; Lumex Inc, Ronkonkoma, NY 11779-0903.
b. Version 11.0; SPSS Inc, 233 S Wacker Dr, 11th Fl, Chicago, IL 60606.
Symmetry of Timing of Hip and Lumbopelvic Rotation Motion in 2 Different Subgroups of People With Low Back Pain

Linda R. Van Dillen, PhD, PT, Sara P. Gomboatto, MSPT, Dave R. Collins, PhD, Jack R. Engsberg, PhD, Shirley A. Sahrmann, PhD, PT


OBJECTIVES: To examine whether lumbopelvic motion associated with a clinical test of active hip lateral rotation (HLR) systematically varied between people classified into 1 of 2 low back pain (LBP) subgroups: lumbar rotation (Rot) or lumbar rotation with extension (RotExt); and, specifically, to determine whether the timing of hip and lumbopelvic rotation with HLR would be more symmetric, right versus left, in people in the Rot subgroup compared with the RotExt subgroup.

Design: Two-group, cross-sectional.

Setting: A university-based movement science laboratory.

Participants: Subjects were 39 people (23 men, 16 women; mean age, 28.1 ± 8.0 y) with chronic or recurrent LBP who regularly participated in a rotation-related sport and associated their LBP symptoms with participation.

Interventions: Not applicable.

Main Outcome Measures: Subjects participated in a standardized clinical examination to classify their LBP problem. A 3-dimensional movement system was used to capture kinematics of hip and lumbopelvic rotation during the test of active HLR. To examine timing of motion between the hip and lumbopelvic region, the difference in time between the start of hip and lumbopelvic rotation was calculated (startdiff). Symmetry of motion was indexed by the correlation (r) between right and left startdiff and the coefficient of determination (r²) for each LBP subgroup.

Results: There were no significant differences between the 2 groups with regard to subject, LBP, activity, and range of motion variables (P > .05 for all comparisons). People in the Rot subgroup displayed significantly more symmetry of timing of hip and lumbopelvic rotation motion with active HLR than people in the RotExt subgroup (Rot subgroup: r = .94, r² = .88, P = .00; RotExt subgroup: r = .31, r² = .10, P = .12).

Conclusions: People in the Rot and RotExt subgroups displayed systematic differences in how they moved the hip and lumbopelvic region with the clinical test of active HLR. These findings are potentially important because such differences in movement patterns between subgroups of people with LBP suggest different contributing factors and may require different treatments to affect the movement patterns.

Key Words: Classification; Hip; Low back pain; Rehabilitation; Rotation.

© 2007 by the American Congress of Rehabilitation Medicine and the American Academy of Physical Medicine and Rehabilitation

Because of the anatomic proximity and interconnections of the hip joint and lumbopelvic region, a number of investigators have focused on the relationship between hip mobility and low back pain (LBP). The interest in the hip-LBP relationship is based on the proposal that limited hip motion will be compensated for by motion in the lumbopelvic region. The proposed result is (1) an increase in the frequency of lumbopelvic motion with hip motion, (2) low magnitude loading in the lumbar region, (3) accumulation of tissue stress, and eventually (4) LBP symptoms.

A number of investigators have focused on the relationship between hip rotation motion and LBP.1-11 Thus far, the primary movement characteristic examined has been end range active or passive hip rotation range of motion (ROM). Based on these studies some people with LBP appear to have less active or passive hip rotation when compared with people without LBP.1,3-5 and subgroups of people with LBP may display different patterns of end range hip rotation motion.3,6,12-14 Although the findings from these studies suggest that hip rotation mobility and LBP may be related in some people with LBP, the nature of the relationship is still not fully understood.

Two issues, in particular, may contribute to the difficulty understanding the relationship between hip rotation and LBP. One issue is that in the majority of prior studies, the variable measured is end range hip motion. A focus on end range hip motion in people with LBP assumes that lumbopelvic motion occurs with hip rotation and occurs at the end of the available hip rotation motion. Such a focus does not consider that most functional activities are performed in the early and mid-ranges of joint movement instead of at the extreme of ROM. To our knowledge, no prior studies have actually quantified the magnitude and timing of lumbopelvic rotation during hip lateral rotation motion, or characterized the symmetry, right versus left, of the timing of hip and lumbopelvic rotation. Examination of the timing of hip and lumbopelvic rotation movement characteristics may provide more insight into the potential relationship between hip rotation motion and LBP than has been evident thus far.

A second issue is that most of the prior studies have been conducted on undifferentiated groups of people with LBP. There are data to suggest that people with LBP can be classified into subgroups based on information obtained during a clinical examination.15-18 Differentiating people with LBP into more homogeneous subgroups may enhance the ability to detect
differences in movement characteristics between subgroups during movement tests such as hip rotation.

In particular, Sahrmann et al.\textsuperscript{16,19} have described the Movement System Impairment (MSI) classification system for classifying people with LBP into subgroups based on symptoms reported and patterns of movement and alignment identified during a standardized clinical examination. The proposed subgroups are named for the directions of movement and alignment that appear to be related to the LBP problem. The subgroups include lumbar rotation (Rot), lumbar flexion, lumbar extension, lumbar rotation with flexion (RotFlex), and lumbar rotation with extension (RotExt). In contrast to other examinations for people with LBP, tests of active limb movements are included. These tests are included because of the potential impact of limb movements on the lumbopelvic region and on LBP symptoms. In particular, the examination includes a test of active hip lateral rotation (HLR) performed in prone. During the HLR test, LBP symptoms are monitored and a judgment is made of whether lumbopelvic rotation occurs in the first 50% of the HLR movement.

One movement characteristic that has been a focus of studies of people with LBP has been the degree of symmetry, right versus left, of end range trunk movement.\textsuperscript{20,21} Some investigators have suggested that there are subgroups of people with LBP who can be identified based on whether they display symmetrical or asymmetrical movement during trunk movement tests.\textsuperscript{16,17,22} Currently, the primary movement test in which symmetry of end range trunk motion is judged is trunk lateral bending in standing. Sahrmann,\textsuperscript{18} however, has proposed that symmetry of trunk motion should be assessed not only with tests of overt trunk movements but also with tests of active limb movements. The proposal is based on clinical observation that some people with LBP appear to move symmetrically with various trunk and limb movement tests and with functional activities, whereas other people with LBP appear to move asymetrically. In particular, it appears that people in the Rot subgroup move symmetrically whereas people in the RotExt LBP subgroup move asymmetrically. Although we have identified movement of the lumbopelvic region during the clinical test of HLR, the data is based on operationally defined criteria used for clinician judgment, and has been limited to the categorical level (present or absent). Kinematic analysis of movements during the clinical test of HLR would allow movement characteristics during the test to be captured with greater precision than the operationally defined responses made at the clinical level.

The purpose of the current study, therefore, was to examine the symmetry of timing of lumbopelvic rotation relative to hip rotation, right versus left, during a clinical test of active HLR using instrumented measures. Specifically, we compared people in the Rot subgroup with people in the RotExt subgroup. We hypothesized that people in the Rot subgroup would exhibit more symmetry of timing of hip and lumbopelvic rotation, right versus left, during a clinical test of active HLR compared with people in the RotExt subgroup. Examination of how people in different LBP subgroups move during a clinical movement test is potentially important because the findings may provide (1) insight into the movement patterns that a particular subgroup has adopted that may be contributing to the LBP problem, and (2) a basis for more specific prevention and treatment strategies.

METHODS

Participants

A convenience sample of 39 people (23 men, 16 women; mean age, 28.1 ± 8.0y), with chronic or recurrent LBP, participated in the study. The sample was a subset of people with LBP who had been recruited from the community for a larger study. The primary focus of the larger study was on the relationship between participation in rotation-related activities and types of impairments displayed. All subjects were people who reported a history of LBP for a minimum of 1 year. Subjects also reported (1) regular participation (minimum 2 times a week) in a sport that placed repetitive rotational demands on the hip and lumbopelvic region, and (2) an increase in their LBP symptoms related to their sport that occurred either during or after play. Examples of activities that were considered to place rotational demands on the hip and lumbar region included sports such as tennis, squash, racquetball, and golf. No subject was in an acute flare-up\textsuperscript{23} of an LBP problem at the time of testing. A flare-up in the current report is defined as a phase of pain superimposed on a recurrent or chronic course which consists of a period, usually a week or less, when the LBP is markedly more severe than usual for the patient. Using a forced choice format, people were screened for participation through a telephone survey and then with a list of questions asked on the day of testing. We excluded people from participation if they reported that a physician had diagnosed them with any of the following spine-related conditions: (1) previous spinal surgery, (2) marked kyphosis or scoliosis, (3) spondylolisthesis, (4) spinal stenosis, (5) spinal instability, (6) spinal fracture, (7) ankylosing spondylitis, (8) degenerative disk disease, (9) disk herniation, or (10) serious spinal complications (eg, tumor or infection). People were also excluded if they had been diagnosed by a physician with (1) rheumatoid arthritis, (2) any condition resulting in severe neurologic involvement, (3) neurologic disease which required hospitalization, (4) history of unresolved cancer, (5) osteoporosis, or (6) were currently pregnant. Finally, people who reported any lower-extremity impairment, for example, impairment secondary to a prior surgery or leg-length discrepancy, were also excluded. Subjects read and signed an informed consent statement approved by the Washington University School of Medicine Human Studies Committee before participating in the study.

Procedures

Self-report measures. Subjects first completed 4 self-report measures. One measure was a numeric rating scale which consisted of an 11-point scale (range, 0–10) provided verbally to the person. The person rated his current symptoms in standing and his average and worst symptoms over the prior 7 days, with higher numbers indicating higher symptom intensity.\textsuperscript{24} A second measure was the Oswestry Disability Index (ODI),\textsuperscript{25} a measure of perceived LBP-related disability. The total score on the ODI can range from 0% to 100% (0% indicating no disability). A third measure was the Baecke Habitual Activity questionnaire,\textsuperscript{26} which is a 16-item questionnaire that measures habitual physical activity over a prolonged period of time. Finally, a fourth measure was a questionnaire of demographic and LBP history variables recommended for reports of studies involving people with LBP.\textsuperscript{27} Summary data related to self-report measures are presented in table 1.

Laboratory testing. Subjects then participated in laboratory testing. We used a 3-dimensional, 6-camera, motion measurement system\textsuperscript{a} to examine the kinematics during the HLR test. The motion measurement system has a sampling rate of 60Hz for each camera and a static resolution of 1mm for a volume of 1m\textsuperscript{3}. Movement of the lower extremity and lumbopelvic region were captured during the HLR test through the use of 20 retro-reflective markers placed on pre-determined

...
Symptoms with the HLR test were obtained. Symptoms with extremity. No subject exceeded the 10 second limit. Reports of speed. A 10-second period was allotted to complete the hip position. Hip movements were performed at a self-selected midline as far as possible, and then return it to the starting laterally rotate their hip so to bring their foot in toward the
sively positioned in 90° of knee flexion and neutral hip abduc-
performed in prone. The subject’s lower extremity was pas-
in the larger study, including the HLR test. The HLR test was
appendix 1, fig 1).

Table 1: Differences in Characteristics of Subjects in 2 LBP Subgroups

<table>
<thead>
<tr>
<th>Variable</th>
<th>Lumbar Rotation (n=13)</th>
<th>Lumbar Rotation With Extension (n=26)</th>
<th>Statistical and Probability Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y)</td>
<td>29.0±7.9</td>
<td>27.7±8.2</td>
<td>t26 = -0.47, P = .64</td>
</tr>
<tr>
<td>Sex (male/female)</td>
<td>6/7</td>
<td>17/9</td>
<td>χ² test = 1.33, P = .25</td>
</tr>
<tr>
<td>Body mass index (kg/m²)</td>
<td>25.5±5.5</td>
<td>24.4±2.5</td>
<td>t26 = -0.88, P = .39</td>
</tr>
<tr>
<td>Right handed (%)</td>
<td>92</td>
<td>96</td>
<td>χ² test = .26, P = .61</td>
</tr>
<tr>
<td>Positive for magnified symptom behavior22 (%)</td>
<td>0</td>
<td>0</td>
<td>Statistical test not performed due to lack of variability in data.</td>
</tr>
<tr>
<td>Positive for neurologic involvement (%)</td>
<td>0</td>
<td>0</td>
<td>Statistical test not performed due to lack of variability in data.</td>
</tr>
<tr>
<td>Symptom intensity* (0–10)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current in standing</td>
<td>2.3±1.9</td>
<td>2.3±1.6</td>
<td>t26 = -10, P = .92</td>
</tr>
<tr>
<td>Average over prior 7 days</td>
<td>3.6±2.0</td>
<td>3.0±1.7</td>
<td>t26 = -1.00, P = .32</td>
</tr>
<tr>
<td>Worst over prior 7 days</td>
<td>5.0±2.0</td>
<td>4.7±2.4</td>
<td>t26 = -0.48, P = .63</td>
</tr>
<tr>
<td>Location of symptoms (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low back only1</td>
<td>100</td>
<td>100</td>
<td>Statistical test not performed due to lack of variability in data.</td>
</tr>
<tr>
<td>Duration of LBP (y)</td>
<td>6.4±5.7</td>
<td>6.1±5.7</td>
<td></td>
</tr>
<tr>
<td>Reproduction of LBP symptoms during hip lateral rotation test (%)</td>
<td>38</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>No increase in LBP</td>
<td>38</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>Right HLR</td>
<td>23</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Left HLR</td>
<td>15</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>Both</td>
<td>23</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>Ability to modify increased symptoms with primary test of hip lateral rotation test during secondary test1 in clinical examination (%)</td>
<td>100</td>
<td>100</td>
<td>Statistical test not performed due to lack of variability in data.</td>
</tr>
<tr>
<td>Episodes of LBP23 (prior 12mo)</td>
<td>14.4±22.0</td>
<td>6.2±8.5</td>
<td>t26 = -1.70, P = .10</td>
</tr>
<tr>
<td>ODI scores25 (0%–100%)</td>
<td>21.2±17.7</td>
<td>13.4±6.1</td>
<td>t26 = -2.03, P = .05*</td>
</tr>
<tr>
<td>Median</td>
<td>18.0</td>
<td>12.0</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>64.0</td>
<td>26.0</td>
<td></td>
</tr>
<tr>
<td>Frequency of participation in rotation-related sport (sessions per week)</td>
<td>3.0±1.9</td>
<td>3.0±1.6</td>
<td>t26 = .00, P = 1.00</td>
</tr>
<tr>
<td>Duration of participation in rotation-related sport (min/session)</td>
<td>92.3±25.9</td>
<td>100.4±60.3</td>
<td>t26 = .46, P = .65</td>
</tr>
<tr>
<td>Baecke activity score26 (3–15)</td>
<td>8.2±0.7</td>
<td>8.4±0.9</td>
<td>t26 = .58, P = .57</td>
</tr>
<tr>
<td>Movement time of HLR angle (s), mean ± standard error</td>
<td>Left = 2.1±0.3</td>
<td>Left = 2.6±0.2</td>
<td>F = .01, P = .93 (subgroup by side)</td>
</tr>
<tr>
<td></td>
<td>Right = 2.1±0.3</td>
<td>Right = 2.6±0.2</td>
<td></td>
</tr>
</tbody>
</table>

NOTE: Values are mean ± standard deviation (SD) or as otherwise indicated.
*Verbal numeric rating of symptom intensity on a scale of 0 (no symptoms present) to 10 (worst imaginable symptoms).24
1Low back region is defined as the area between T12-L1 interspace and the gluteal fold.22
2Secondary tests are conducted if the patient reports an increase in symptoms with a primary test. Modifications for movement tests involve restricting or eliminating lumbar region movement during a trunk or limb movement while encouraging movement in other segments, for example, the hips or thoracic region. Modifications are accomplished using verbal cues, trunk muscle activation by the person, and manual assistance by the examiner. The person reports his symptoms with each secondary test relative to the symptomatic primary test.
3Indicates a significant difference between the 2 LBP subgroups.

anatomic landmarks of the trunk, pelvis and extremities (appendix 1, fig 1).

A series of laboratory-based movement tests were conducted in the larger study, including the HLR test. The HLR test was performed in prone. The subject’s lower extremity was passively positioned in 90° of knee flexion and neutral hip abduction and adduction and neutral rotation. Subjects were asked to laterally rotate their hip so to bring their foot in toward the midline as far as possible, and then return it to the starting position. Hip movements were performed at a self-selected speed. A 10-second period was allotted to complete the hip movement and the movement was performed once with each extremity. No subject exceeded the 10 second limit. Reports of symptoms with the HLR test were obtained. Symptoms with the HLR test were relative to the subject’s symptoms in prone. The possible options for symptom reports included the following: (1) remained the same, (2) increased, (3) decreased, or (4) eliminated.

**Laboratory Data Processing**

Initially, all marker data were filtered using a fourth-order, dual-pass, Butterworth filter with a cutoff frequency of 2.5Hz. The initial cutoff frequency was chosen because the movements being measured were relatively slow. Because subjects performed the test movement at self-selected speeds, a filtering frequency based on individual movement times was then used. Based on the HLR movement time, raw data were filtered at a
subject-specific cutoff frequency\(^{28}\) (\(f_{c,s}\)) that was calculated by taking the reciprocal of 15% of the period:

\[
f_{c,s} = 1/\left(0.15 \times (2 \times \text{HLR movement time})\right).
\]

Clinical Examination and Classification

Subjects then participated in a standardized clinical examination used to classify the person’s LBP problem.\(^{18,19}\) Testing of the reliability of examiners performing the physical tests and measures from the standardized examination has been reported.\(^{19}\) Fair to good reliability\(^{29}\) of classifying a patient’s LBP has also been reported (\(\kappa = 0.57\); percentage of agreement, 78%)\(^{30}\) by our own research group, and an independent research group (\(\kappa = 0.56\)).\(^{31}\)

The examination includes tests of movements and positions, neurologic screening, and tests to screen for magnified symptom behavior.\(^{32}\) Each test is proposed to be associated with a particular direction of lumbar region movement or alignment, including flexion, extension, rotation, or a combination of rotation and flexion or rotation and extension. Because rotation and lateral bending are coupled motions in the lumbar spine,\(^{33,34}\) symptoms or patterns of movement or alignment identified during examination tests associated with either rotation or lateral bending currently are labeled rotation. Symptoms are monitored with each test, and for some of the tests, judgments of patterns of movement and alignment are made. The judgments are made by the examiner based on visual or on both visual and tactile information. The response options for LBP symptoms are the same as those described for the laboratory testing. With each initial test, referred to as a primary test, the subject assumes a position or performs a movement using a preferred strategy. If the subject reports an increase in symptoms with the primary test, it is immediately followed by a secondary test in which the subject’s preferred strategy is standardly modified.\(^{35,36}\) The secondary test is performed in an attempt to decrease or eliminate the subject’s symptoms. Based on the consistency of findings of lumbar region movement and alignment and LBP symptom behavior across the clinical examination, each subject is classified into 1 of 5 LBP subgroups.\(^{18}\) The subgroups are named for the directions of movement and alignment that appear to contribute to the LBP problem. The subgroups include (1) Rot, (2) lumbar flexion, (3) lumbar extension, (4) RotFlex, and (5) RotExt. In deciding a person’s classification, priority is given to the symptomatic primary tests that can be successfully modified (symptoms decreased or eliminated). Findings typical of the Rot and RotExt LBP subgroups are described in table 2.

Kinematic Measures

**Hip and lumbopelvic rotation.** We calculated hip lateral rotation and lumbopelvic rotation angles across each 10-second trial of HLR. Hip lateral rotation was calculated by tracking the lower leg segment across time. The lower leg segment was defined by a vector from a marker on the lateral aspect of the knee joint line to a marker at the distal aspect of the lateral malleolus. Angle \(\alpha\) was calculated as the change in angle of the lower leg segment across time, relative to its initial position (see fig 1). Assuming that no relative motion occurs between the tibia and femur, movement of the tibia should reflect rotation of the femur. Degree of HLR was then defined by subtracting, from angle \(\alpha\), pelvic motion that occurred in the plane of the lower leg. Thus, movement of the pelvis that resulted in movement of the lower leg was not included in the hip rotation calculation.

Lumbopelvic rotation was calculated by tracking the pelvic segment across time. The pelvic segment was defined as a vector formed by the right and left iliac crests. Relative to the initial position of the pelvic segment, the position of the pelvic vector was calculated across time during the HLR motion to calculate \(\theta\), the degree of lumbopelvic rotation (see fig 1).

The criteria that we used to identify the optimal start and end of movement for hip and lumbopelvic rotation were determined through an iterative process. The hip lateral rotation angle and lumbopelvic rotation angle for each subject was plotted against time. Start and end points for each angle were identified using threshold criteria of (1) angular displacement and (2) angular velocity. The threshold values were varied and the start and end of both angles, for each subject, were plotted and visually inspected. For visual inspection, an accurate start point was defined as the time point at which there was a consistent change in the slope of the angle-time plot. An accurate end point was defined as the first time point at which the angle-time plot appeared to reach the absolute maximum. The threshold values for angular displacement and velocity that resulted in the most accurate start and end points for the majority of subjects were used in the final algorithm. The start of movement for HLR was defined as the time at which both (1) angular displacement of the lower leg segment exceeded a threshold of 1° and (2) angular velocity exceeded 5% of the maximum angular velocity for the lower leg segment. The start of movement for lumbopelvic rotation was defined as the time at which both (1) the angular displacement of the pelvic segment exceeded a threshold of 1° and (2) the angular velocity exceeded 15% of the maximum angular velocity for the pelvic segment. The end of movement for the HLR and lumbopelvic rotation angles were defined by the first point at which each angle reached 99% of its maximum during the hip rotation movement.

We tested the reliability of HLR and lumbopelvic rotation measures in a sample of 10 subjects without a history of LBP. The intraclass correlation coefficient model 3,1 (ICC\(_{3,1}\))\(^{37}\) and standard error of the measure\(^{38,39}\) were used to index reliability. The values for each motion, for each extremity, were found acceptable and are reported in table 3.

Data Analyses

**Dependent variables.** Select descriptive variables related to subject characteristics, self-report, examination, and labora-

---

Fig 1. Kinematic model indicating angles for HLR and lumbopelvic rotation. Abbreviations: \(\alpha\), HLR angle (in degrees); GT, greater trochanter marker; IC, iliac crest marker; K, knee marker; LM, lateral malleolus marker; \(\theta\), lumbopelvic rotation angle (in degrees). From Gombatto et al.\(^{32}\) Reprinted with permission.
Statistical analyses. Data analysis was performed using Systat for Windows. Frequency counts and percentages of the subjects in the 2 LBP subgroups were calculated. A chi-square analysis was conducted on the distribution of sex across the 2 LBP subgroups (Rot, RotExt). Descriptive statistics and tests of differences between the 2 subgroups were calculated on select subject characteristics, self-report, examination, and laboratory variables.

Table 3: ICC Values and Standard Error Measures and for Right and Left HLR and Lumbopelvic Rotation

<table>
<thead>
<tr>
<th>Kinematic Measure</th>
<th>ICC&lt;sub&gt;3,1&lt;/sub&gt;</th>
<th>Standard Error of Measure (deg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left HLR angle</td>
<td>.99</td>
<td>1.5</td>
</tr>
<tr>
<td>Right HLR angle</td>
<td>.82</td>
<td>0.8</td>
</tr>
<tr>
<td>Left lumbopelvic rotation angle</td>
<td>.97</td>
<td>1.4</td>
</tr>
<tr>
<td>Right lumbopelvic rotation angle</td>
<td>.80</td>
<td>0.8</td>
</tr>
</tbody>
</table>

From Gombatto et al. Reprinted with permission.

A 2-way mixed-model analysis of variance was conducted to test for main and interaction effects of LBP subgroup (Rot, RotExt) and side (right, left) on maximum hip lateral rotation angle and maximum lumbopelvic rotation angle. Degree of

![Figure 2. Angle of HLR and lumbopelvic rotation as a function of time. Abbreviations: LPR, lumbopelvic rotation; startdiff, difference in time between the start of HLR and lumbopelvic rotation.](image)
symmetry in timing of HLR and lumbopelvic rotation was indexed by 2 related statistics. The first statistic calculated was the Pearson product-moment correlation coefficient that indexes the linear relationship between 2 variables. The correlation coefficient was calculated between the startdiff variable for right HLR and the startdiff variable for left HLR. Symmetry of timing, right versus left, would be represented by a coefficient approaching 1.0, whereas asymmetry of timing would be represented by a coefficient approaching zero. The correlation coefficient was examined for each LBP subgroup (Rot, RotExt). Differences in symmetry of timing of HLR and lumbopelvic rotation between subgroups was tested by converting the correlation coefficient for each subgroup to z scores, estimating the standard error of the difference based on sample size, and testing for a significant difference in z scores between subgroups. The second statistic calculated was the coefficient of determination or $r^2$ value. The $r^2$ value represents the amount of shared variance between 2 variables and denotes the strength of the linear association, that is, proportion of shared variance between 2 variables. The coefficient of determination was calculated for each LBP subgroup. All significance testing was set at the $P$ equal to or less than .05 level.

RESULTS

Thirteen (33%) of the subjects were classified as Rot and 26 (67%) of the subjects were classified as RotExt. The proportion of people within the 2 subgroups was significantly different ($\chi^2$ test=4.33, $P$=.04). No subjects displayed signs of neurologic involvement or magnified symptom behavior.32

Table 1 provides the results of the tests of differences between the subgroups with regard to subject-, LBP-, examination-, and movement-related characteristics. The subgroups were not statistically different with regard to any variables ($P$>.05 for all comparisons) except for the ODI score.25 People in the Rot subgroup tended to score higher on the ODI compared with people in the RotExt subgroup ($P$=.05).

Kinematic Variables

We compared the 2 subgroups with regard to maximum HLR and lumbopelvic rotation angle. The maximum angles for each subgroup, for each side, are provided in table 4. There were no differences between subgroups for maximum HLR angle (Rot, 41.9°±1.8°; RotExt, 41.1°±1.3°; $F_1=2.08$, $P$=.16), and there were no differences between sides (left, 43.7°±1.2°; right, 44.3°±1.2°; $F_1=0.3$, $P$=.86). There was also no interaction of subgroup and side for maximum HLR angle (see table 4). There also were no differences between subgroups for maximum lumbopelvic rotation angle (Rot, 4.8°±0.8°; RotExt, 6.0°±0.6°; $F_1=1.42$, $P$=.24), and there were no differences between sides (left, 5.3°±0.5°; right, 5.9°±0.6°; $F_1=.98$, $P$=.33). There was no interaction of subgroup and side for maximum lumbopelvic rotation angle (see table 4).

We then examined the correlation coefficients indexing the relationship between right and left startdiff values for each subgroup and the coefficient of determination for each subgroup. People in the Rot subgroup showed more symmetry in timing of HLR and lumbopelvic rotation, right versus left, than people in the RotExt subgroup. The correlation coefficient indexing the relationship of right and left startdiff for the Rot subgroup was .94 and the $r^2$ value was .88 ($P$=.00) (fig 3A).

### Table 4: Differences of HLR and Lumbopelvic Rotation During Kinematic Testing Between People in the Lumbar Rotation LBP Subgroup Versus People in the Lumbar Rotation With Extension LBP Subgroup

<table>
<thead>
<tr>
<th>Variable</th>
<th>Lumbar Rotation (n=13)</th>
<th>Lumbar Rotation With Extension (n=26)</th>
<th>Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Left</td>
<td>Right</td>
<td>Left</td>
</tr>
<tr>
<td>Active HLR (deg)</td>
<td>42.3±2.0</td>
<td>41.5±2.1</td>
<td>44.5±1.5</td>
</tr>
<tr>
<td>Lumbopelvic rotation (deg)</td>
<td>4.8±0.8</td>
<td>4.9±1.0</td>
<td>5.6±0.6</td>
</tr>
</tbody>
</table>

NOTE. Values are mean ± SD.

*Indicates the statistical results for the 2-way interaction of subgroup by side. There were also no main effects of LBP subgroup or side.
The correlation coefficient for the RotExt subgroup, on the other hand, was .31 and the \( r^2 \) value was .10 (\( P=.12 \)) (fig 3B). The correlation for the Rot subgroup differed significantly from the correlation for the RotExt subgroup (\( \rho=3.74, P<.01 \)).

**DISCUSSION**

The purpose of the current study was to examine the symmetry of timing of hip and lumbopelvic rotation, right versus left, during a clinical test of active HLR. We compared people who were assigned to 1 of 2 LBP subgroups, the Rot LBP subgroup or the RotExt LBP subgroup. The findings from the current study support the hypothesis that people in the Rot subgroup will show more symmetry, right versus left, in timing of rotation of the hip and lumbopelvic region when compared with people in the RotExt subgroup. More specifically, the timing of hip and lumbopelvic region movement with HLR was essentially the same with the right and left extremity for the people in the Rot subgroup. People in the RotExt subgroup, on the other hand, displayed differences in the timing of hip and lumbopelvic motion with right and left HLR. These findings are important because they provide evidence that specific subgroups of people with LBP show systematic differences in how they move their hip and lumbopelvic region during a standardized clinical movement test. These findings are also important because, unlike prior studies, they are the first to quantify the timing of lumbopelvic rotation during HLR and the differences in timing, right versus left, of hip and lumbopelvic rotation movements between LBP subgroups. Such differences in timing of movement between subgroups during a movement test are of clinical significance because these findings suggest potential differences in movement of the hip and lumbopelvic region during more functional movements. These findings also raise the issue of whether different factors, such as biomechanic or motor control variables, could contribute to the different movement patterns observed in the 2 LBP subgroups. Last, such differences in symmetry of timing, right to left, suggest a need for differences in the approach to intervention that addresses the movement patterns characterizing the different LBP subgroups.

**Model for Development of LBP**

The MSI classification system for LBP is based on the kinesiophore model of movement (KPM),\(^{30}\) a conceptual framework for the mechanisms proposed to contribute to the development of musculoskeletal pain. The KPM, therefore, provides a theoretical basis for treatment of musculoskeletal pain problems, including LBP. The basic premise of the KPM is that musculoskeletal pain develops as a result of loss of movement precision, that is, an alteration in appropriate biomechanics, at 1 or more segments in a specific anatomic region. Loss of movement precision is proposed to develop when movements are repeated and postures are sustained in a particular direction(s) during everyday activities, inducing changes in neural and musculoskeletal components of the movement system. Such changes then result in adoption of direction-specific strategies of movement and alignment that become generalized across many activities. The specific loss of precision proposed for development of LBP is an increase in flexibility of 1 or more lumbar segments in a specific direction(s), with a potential decrease in flexibility in other regions, for example the thoracic region or hips. The low magnitude loading with repetition in the same direction is considered to contribute to acceleration of accumulation of tissue stress in the lumbar region due to minimal time without loading,\(^{31}\) microtrauma, and eventually LBP symptoms. The model further contends that until the specific precision loss is addressed, the symptoms have the potential to persist or recur. The subgroups proposed in the MSI classification system display the most prevalent directions of movement and alignment strategies that are proposed to characterize LBP problems.

**Treatment Based on KPM**

Based on the KPM, identifying a person’s LBP subgroup provides a basis for treatment. Specifically, to improve a person’s LBP problem, treatment would be directed at correcting the direction-related movements and alignments that characterize the person’s particular LBP subgroup. The RotExt subgroup has been described as characterized by a tendency to move and align the lumbar region in rotation and extension across several examination tests and with functional activities. The rotation-related movements and alignments are also often asymmetric\(^{18}\) (see table 2). In general, asymmetry has been proposed to be a potential mechanism for LBP because of the increased frequency of loading of one or more lumbar segments on one side compared with the other.\(^{32}\) In the current study the RotExt subgroup displayed an asymmetry in timing of hip and lumbopelvic region rotation, right to left, with the HLR test. In particular, lumbopelvic rotation occurred early in the range of hip lateral rotation in 1 lower extremity compared with the other (see fig 3B). Specific to the HLR test findings, the precision loss is a relative increase in lumbopelvic rotation flexibility compared with hip rotation flexibility evidenced by lumbopelvic rotation early in the range of HLR motion with 1 lower extremity compared with the other. Treatment would be directed at the neural and musculoskeletal components that appear to contribute to the asymmetry with the HLR test, as well as with other examination tests and functional activities. The Rot subgroup, on the other hand, has been described as characterized by a tendency to move and align the lumbar region in rotation across several examination tests and with functional activities\(^{18}\) (see table 2). The potential mechanism for LBP is the increased frequency of loading lumbar region segments in rotation due to adoption of rotation-related strategies across many activities. The Rot subgroup in the current study displayed symmetry in the timing of hip and lumbopelvic rotation, with the majority of subjects displaying lumbopelvic rotation early in the range of hip lateral rotation motion with both the right and left lower extremities (see fig 3A). Specific to the HLR test findings, the specific precision loss is the increase in lumbopelvic region rotation flexibility relative to hip rotation flexibility bilaterally for the majority of subjects. Treatment would be directed at the neural and musculoskeletal components that appear to contribute to the repeated lumbopelvic region rotation with the HLR test, as well as similar patterns of lumbopelvic rotation with other examination tests and functional activities.

**ODI Scores**

The Rot subgroup had a slightly higher mean ODI score (\( P=.05 \)) (see table 1) than the RotExt subgroup. The difference in LBP-related disability could pose an alternative explanation for the differences in symmetry of timing of hip and lumbopelvic region rotation between the 2 subgroups. The difference in ODI scores, however, was found to be the result of one case in the Rot subgroup that was identified as an outlier\(^{16}\) (ODI score, 68%). Similar to other subjects in the Rot subgroup, however, the timing of lumbopelvic and HLR movement, right versus left, for the outlier case was similar (right startdiff, .22s; left startdiff, .20s). Thus, it does not appear that a difference in
LBP-related disability is the reason for the differences between the 2 LBP subgroups in symmetry of timing.

Prior Studies

Some earlier studies have examined the relationship between active hip lateral rotation and LBP. The variable examined in each study, however, has been the amount of end range hip lateral rotation motion and not the timing of movement of the hip and lumbopelvic region during hip lateral rotation. For example, Mellin reported no relationship between the amount of end range active hip lateral rotation and reported baseline symptom levels in people with chronic or recurrent LBP. Mellin also reported no differences in amount of end range active hip lateral rotation between young adults and with and without a history of LBP. Chesworth et al, on the other hand, reported decreased end range active hip lateral rotation in people with LBP compared with matched subjects without a history of LBP. The current study differs from these prior studies in that (1) both active hip lateral rotation and lumbopelvic rotation were measured, (2) different subgroups of people with LBP were compared, and (3) relative timing of hip lateral rotation and lumbopelvic rotation between sides was the variable of interest rather than end range hip lateral rotation. Because of these differences, the current findings regarding active hip lateral rotation cannot be directly compared with prior studies. In the current study, however, we focused on the relative timing of active hip lateral rotation and lumbopelvic rotation between sides for a number of reasons. First, our clinical observation and prior data suggest that a number of people with LBP report an increase in LBP with the active HLR test. Second, we have reported that some people with LBP also display early lumbopelvic rotation with HLR. The lumbopelvic rotation that occurs with active HLR appears to be of importance because restricting lumbopelvic rotation during the HLR test results in a decrease or elimination in symptoms in the majority of people with LBP. Fourth, the prevalence of lumbopelvic rotation during active hip lateral rotation appears to be related to the types of activities in which people with LBP participate regularly. Specifically, it was found that a larger proportion of people with LBP who participated in asymmetric activities showed asymmetry of lumbopelvic rotation with the HLR test than people with LBP who participated in symmetric activities.

Some of the prior studies have also examined the relationship between end range hip medial rotation and LBP. The findings from these studies suggest that in various LBP subgroups, decreased passive or active end range hip medial rotation or a decrease in end-range passive medial rotation compared with lateral rotation in 1 extremity may be related to the LBP problem. Finally, a discrepancy in the amount of passive hip medial rotation between sides has been associated with a particular type of LBP that appears to respond positively to spinal manipulation. We have documented that people with LBP report an increase in LBP symptoms with the test of active hip medial rotation. Similar to active hip lateral rotation, we have observed that a factor that appears to contribute to the increase in symptoms with the active hip medial rotation test is the attempt to move the lumbopelvic region during the hip motion. At the time of the current study, however, we did not have data regarding the pattern of lumbopelvic rotation with hip internal rotation. Future studies could focus on quantifying, with clinical and instrumented measures, the amount and timing of lumbopelvic motion that occurs with active hip medial rotation in people with LBP, and its relationship to LBP symptoms.

Symmetry of Movement

Trunk lateral bending in standing is the test movement in which symmetry of trunk movement is most often examined. Asymmetry of trunk lateral bending is considered to be an important finding in determining the classification of a person’s LBP problem. In contrast to prior work, we have examined symmetry of trunk movement, indexed by lumbopelvic movement, with an active limb movement test. The current data suggests that a limb movement, HLR, can induce movement of the lumbopelvic region, and that people in the RotExt subgroup display more asymmetry of lumbopelvic rotation than people in the Rot subgroup, at least with hip lateral rotation. Future studies could examine (1) the symmetry of lumbopelvic movement with other limb movement tests in people in the Rot and RotExt LBP subgroups, (2) the relationship of symmetry with trunk movement tests to symmetry with various limb movement tests, and (3) the relationship of symmetry with movement tests to symmetry with more functional movements such as gait. Findings from such studies could provide additional information on which to base treatment of people in the Rot and RotExt LBP subgroups.

Prevalence of LBP Subgroups

The LBP subgroups of interest in the current study are only 2 of 5 subgroups described in the MSI classification system for LBP. The prevalence of the different LBP subgroups has been proposed to differ, with the majority of people described as having an extension, rotation, or combined rotation and extension LBP problem. Two prior studies provide partial support for the described distribution of subgroups. A factor analysis of examination data from 188 people with LBP resulted in the largest proportion of variance explained by tests related to the RotExt subgroup, followed by tests related to the Rot and the Ext subgroup. A subsequent study of 51 people with LBP examined on their first physical therapy clinic visit resulted in the following distribution of LBP subgroups: (1) RotExt (41%), (2) RotFlex (22%), (3) lumbar extension (18%), (4) Rot (14%), and (5) lumbar flexion (5%). In the sample (N=43) considered for the current study, 13 (30%) of the subjects were classified as Rot, 26 (61%) were classified as RotExt, and 4 (9%) were classified as RotFlex. No subjects were classified into the lumbar flexion or extension subgroup. Thus, the distribution of the current sample is similar to our prior studies. We did not include the RotFlex subgroup in the current analyses, however, because (1) our hypothesis was focused specifically on the Rot and RotExt subgroup findings and (2) of the limited number of people in the RotFlex subgroup.

Study Limitations

The current study has some potential limitations. First, we examined the symmetry of hip and lumbopelvic movement during a standardized clinical movement test. Whether these findings generalize to functional activities is not known at this time. Future studies could explore the relationship between movement characteristics identified during standardized movement tests and movement characteristics with functional activities. Second, the findings from our study may not generalize to all people with LBP. The people in the current study were people with LBP recruited from the community who regularly participated in a rotation-related sport. Our findings, therefore, may not be applicable to all people referred to physical therapy clinics or to people who do not regularly participate in rotation-related activities. We do know from prior studies, however, that some subjects recruited from clinics who reported participation in a variety of activities could be classified into the Rot...
or RotExt LBP subgroups. 30,44 Finally, it is possible that the movement patterns we have identified do not contribute to the person’s LBP and may instead be the result of having LBP symptoms. If this were true, intervention directed at the movement patterns characterized an LBP subgroup are modified during the clinical examination and with treatment, symptoms improve for the majority of people with LBP. 35,36,51 We also know from clinical examination and with treatment, symptoms improve that characterize an LBP subgroup are modified during the movement patterns we have identified do not contribute to the movement pattern during HLR was modified to restrict lumbar pelvic rotation (see table 1). Such findings suggest that how subjects were moving during the HLR test appears to be relevant to their LBP symptoms.

CONCLUSIONS

The findings from the current study suggest that people in 2 different LBP subgroups, Rot and RotExt, move differently, right versus left, during the HLR movement test. Specifically, people in the Rot subgroup move the hip and the lumbar pelvic region symmetrically during the HLR test. On the other hand, people in the RotExt subgroup move the hip and lumbar pelvic region asymmetrically during the HLR test. These findings suggest that the symmetry of movement, right versus left, during the HLR test may be an important factor for clinicians to consider when specifying the details of interventions for people in these 2 LBP subgroups.

Acknowledgments: We thank Kevin Hollander, PhD, for his assistance in development of the model used to derive the kinematic measures.

APPENDIX 1: LOCATIONS OF MARKERS FOR CAPTURING MOVEMENT DATA DURING THE HLR TEST

<table>
<thead>
<tr>
<th>Marker Locations*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acromion</td>
</tr>
<tr>
<td>Iliac crest</td>
</tr>
<tr>
<td>Greater trochanter</td>
</tr>
<tr>
<td>Lateral knee joint line</td>
</tr>
<tr>
<td>Distal aspect of lateral malleolus</td>
</tr>
<tr>
<td>Posterior aspect of calcaneus</td>
</tr>
<tr>
<td>C7 spinous process</td>
</tr>
<tr>
<td>T4 spinous process</td>
</tr>
<tr>
<td>T7 spinous process</td>
</tr>
<tr>
<td>T10 spinous process</td>
</tr>
<tr>
<td>L1 spinous process</td>
</tr>
<tr>
<td>L3 spinous process</td>
</tr>
<tr>
<td>L5 spinous process</td>
</tr>
<tr>
<td>S2 spinous process</td>
</tr>
</tbody>
</table>

*All markers are placed bilaterally except spine markers which are placed superficial to specific spinous processes.

References


Suppliers
a. Motion Analysis Corp, 3617 Westwind Blvd, Santa Rosa, CA 95403.
b. Version 10.2; Systat Software Inc, 501 Canal Blvd, Ste E, Point Richmond, CA 94804.
Recovery of Gait After Short-Stay Total Hip Arthroplasty

Inge van den Akker-Scheek, MSc, Martin Stevens, PhD, Sjoerd K. Bulstra, PhD, MD, Johan W. Groothoff, PhD, Jim R. van Horn, PhD, MD, Wiebren Zijlstra, PhD


Objective: To describe recovery of gait after total hip arthroplasty (THA) based on the assessment of spatiotemporal gait parameters determined with an ambulatory system.

Design: A 6-month inception cohort study.

Setting: Inpatient and outpatient setting in an academic hospital.

Participants: Sixty-three patients participating in a short-stay program for THA.

Intervention: Primary unilateral THA.

Main Outcome Measures: Walking speed, step length, step duration, and variability coefficient assessed at different walking speeds while performing an additional cognitive task and an endurance test. All measures were obtained preoperatively and 6 weeks and 6 months postoperatively.

Results: Patients improved significantly over time; however, extent and speed of recovery of gait parameters differed for each test part. The relation between walking speed and step length showed systematic improvement when analyzed over a range of speeds. At 6 months, the variability coefficient of the additional task test part was comparable with the preferred walking variability coefficient. The endurance test results could be predicted from the results of preferred walking.

Conclusions: Assessment of recovery of gait function requires more than only assessment of "normal" walking. Particularly, an analysis of walking at different speeds and walking while performing an additional cognitive task demonstrate different aspects of gait recovery after THA.

Key Words: Arthroplasty, replacement, hip; Gait; Rehabilitation.

© 2007 by the American Congress of Rehabilitation Medicine and the American Academy of Physical Medicine and Rehabilitation

BESIDES SEVERAL WIDELY used self-report questionnaires to describe recovery after total hip arthroplasty (THA), objective measures can be obtained and are considered complementary. In particular, measures of gait function are relevant to assess outcome. Because gait is highly important in everyday life, gait function is closely linked to overall functioning. For example, walking speed has been related to independent living and the ability to perform various activities of daily living, like safely crossing a traffic intersection and risk of falling. Hence, gait function is an important indicator of functional recovery. Gait function depends on the ability to maintain safe gait while navigating in a complex and changing environment, whereby gait has to be adjusted to different task demands. Thus, from a behavioral approach, important elements of gait function are the ability to walk at different speeds, combine gait with attention-demanding tasks, and walk longer distances. These aspects should be incorporated into a test battery assessing THA patients. Next to questionnaire research, predominantly limited gait analyses are performed (ie, only "normal" walking) on THA patients. With the development of ambulatory measuring devices that use body-fixed sensors, gait analysis is less time-consuming compared with the methods applied in laboratories by using camera systems, forceplates, treadmills, and electromyography. It is therefore possible to measure different aspects of gait function within an acceptable amount of time and burden. Additionally, gait analysis can be performed outside a laboratory setting, which is a more representative environment of the patients' everyday life.

The purpose of this study is to describe the spatiotemporal measures of gait recovery after THA, whereby the different aspects of gait function are measured with an ambulatory measuring device in an extensive test protocol, including walking at 3 speeds, walking with an additional cognitive task, and a walking endurance test. The hypothesis is that this test protocol will provide measures of gait recovery that cannot be detected when only preferred walking speed is assessed.

METHODS

Participants

From September 2002 to August 2004, patients were included in the study on admission to the Orthopedic Department of the University Medical Center Groningen for a unilateral primary THA while participating in the short-stay program. Because not all patients are capable of following a short-stay program, selection was done by using the following criteria: estimated surgery time less than 120 minutes, body mass lower than 110kg, expected hospital stay less than 6 days, absence of relevant cardiovascular comorbidity, no severe mobility disability, no psychologic dysfunction, and no severe deformity of the spine. All patients were allowed to start walking with aids on the first postoperative day and were discharged on the fifth postoperative day, unless there were complications. No physical therapy was prescribed.

Within the short-stay program, 2 groups could be distinguished. One group received usual short-stay care, another received an additional support program called the Groningen Orthopedic Exit Strategy (GOES). The GOES was developed to respond to the insecurity and inactivity many patients show during the rehabilitation period, which as a consequence of the short-stay program is moved from the hospital setting to the
home situation and aims to provide patients with extra support. The GOES treatment consists of showing patients a video with role models, providing extra information through newsletters and telephone consultation hours, and emphasizing partner participation. The hypothesis that this treatment would lead to more self-efficacy and social support and better pain coping, and eventually result in a better and quicker rehabilitation, was not supported, as measured by questionnaires. Because no effect on the psychosocial variables was found, it is assumed that no differences in gait recovery between these 2 patient groups will be present and the 2 groups can be evaluated as a whole.

Design and Data Collection
All patients included in the study performed the gait analysis on admission to the hospital preoperatively and 6 weeks and 6 months postoperatively when visiting the outpatient clinic. The greatest incremental improvements in gait parameters have been shown to occur in the first 6 months after surgery. Additionally, a reference population consisting of relatives of the THA patients performed the gait analysis in the same way the patients did. The study was conducted in accordance with the guidelines stated in the Declaration of Helsinki and the regulations of the Medical Ethical Committee of the University Medical Center Groningen.

Measurements
For gait analysis, we used the DynaPort system, which consists of a data recorder (dimensions, 125×95×34mm; weight, 295g) that is attached to the lower back of the patient with a neoprene belt around the waist. The data recorder contains 3 uniaxial, piezoresistive accelerometers that measure acceleration in the frontal, sagittal, and transverse planes and a memory card on which data are stored. Three penlight batteries are attached to the belt. The position of the data recorder is adjustable so that the sensors can be positioned vertically.

Measurements were performed in a normal hospital corridor under supervision of the first author. Two cones were positioned 20m apart. Patients were instructed to walk 20m at their preferred speed, to walk as fast as possible (not running), and to walk at a slow speed. Additionally, patients were instructed to walk at their preferred speed while counting backward out loud from 50, each time subtracting 3 (dual task). Finally, patients were instructed to walk at their preferred speed, to walk as fast as possible (not running), and were unable to walk, 3 patients had health problems unknown reasons, 1 patient moved out of the country, and 2 patients underwent an arthroplasty of the contralateral hip before the final measurements took place. These 10 patients were therefore excluded: 2 patients had severe complications and eventually result in a better and quicker rehabilitation, was not supported, as measured by questionnaires. Because no effect on the psychosocial variables was found, it is assumed that no differences in gait recovery between these 2 patient groups will be present and the 2 groups can be evaluated as a whole.

RESULTS

Demographic Characteristics
A total of 80 THA patients participated in the short-stay program during the inclusion period of the study. Ten of them could not perform the second or third measurement or both and were therefore excluded: 2 patients had severe complications and were unable to walk, 3 patients had health problems unrelated to the arthroplasty at the time of the third measurement, 2 patients did not show up at the outpatient clinic for unknown reasons, 1 patient moved out of the country, and 2 patients underwent an arthroplasty of the contralateral hip before the final measurements took place. These 10 patients had a mean age of 62.2 years, and 70% were women. Body mass index was 29.2kg/m². Six other patients had to be excluded because of technical problems and 1 patient because of the fact that not all test parts of 1 measurement were performed. The data of 63 patients were therefore available for analysis. One person from the reference group had to be excluded because of technical problems so the data of 19 persons were used in the analysis.

Table 1 presents baseline characteristics of the patients and the reference population. The patients’ mean age was 62.0 years, and most (68.3%) patients were women. There were no significant differences between the patients and the reference group on any variable. Mean hospital length of stay (LOS) was 7.0±3.3 days. The mean LOS was raised by 1 patient with an LOS of 26 days because of an infection and prolonged wound leakage. Not including this particular patient, the mean LOS was 6.7±2.3 days.
Gait Parameters

Gait characteristics for the THA patients as well as for the reference population are displayed in table 2. All main effects for measurement time were significant with a P value lower or equal to .001, meaning that on all gait parameters and on all test parts, the patients improved significantly over time.

Between the preoperative and 6-week postoperative measurements, walking speed improved significantly only in the slow-walking and the dual-task test parts, whereas step length and the variability coefficient improved significantly in the slow-walking, preferred-walking, and dual-task test parts. Step duration did not improve significantly in any test part. Between the 6-week and 6-month postoperative measurements, significant improvement was seen in walking speed, step length, and step duration in all test parts. The variability coefficient improved significantly only in the fast walking and dual-task test parts. The relation between the parameters of step length and the variability coefficient improved significantly in the slow-walking and the dual-task test parts, whereas step length did not improve significantly in any test part. Between the preoperative and 6-week postoperative measurements, walking speed improved significantly only in the slow-walking and dual-task test parts, whereas step length and the variability coefficient improved significantly in the slow-walking, preferred-walking, and dual-task test parts. Step duration did not improve significantly in any test part. Between the 6-week and 6-month postoperative measurements, significant improvement was seen in walking speed, step length, and step duration in all test parts. The variability coefficient improved significantly only in the fast walking and dual-task test parts.

Decreasing the variability coefficient improves significantly only in the fast walking and dual-task test parts. The relation between the parameters of step length and walking speed for the 3 measurement times is shown in figure 1.

Between the different test parts preoperatively, most variability was seen in the dual-task test part (fig 2). At 6 weeks, all variability coefficients improved, except in the fast-walking test part. Still, the variability coefficient observed in the dual-task test part was the highest. At 6 months postoperatively, most variability was seen in the slow-walking test part.

Patients improved significantly in the distance they could walk within 6 minutes (F1,7,105.1=43.68, P<.001; partial η2=.41). Preoperatively, 294.0±103.8m were covered on average, 315.9±75.3m at 6 weeks, and 387.4±79.5m after 6 months. On average, the reference group walked 475.6±49.2m. Preoperatively, as well as postoperatively, most patients were able to walk for 6 minutes. The mean values of the first and final three 20-m distances were calculated for all gait parameters at the 3 different measurement times (table 3). Six patients who could not walk at least six 20-m distances were excluded from this analysis. Patients improved significantly over time on the difference between the first and final mean values of the parameters walking speed, step length, and step duration (P<.05); at 6 months, the mean walking speed of the final three 20-m distances was even higher than the mean walking speed of the first 3. Differences between measurement times were very small though; this was expressed in low partial η2 values (.11, .06, .11, respectively). No significant improvement was seen on the difference between the mean first and final three 20-m distances of the variability coefficient.

Of the 63 included patients, 35 received the GOES treatment. The GOES patients were comparable on all patient characteristics measured at baseline to the group of patients receiving usual short-stay care. No significant group by time effect was found on any of the gait parameters of the slow-, preferred-, and fast-walking or dual-task test parts; the patients of the GOES group did not improve significantly over time any differently than those of the usual short-stay group. The walking endurance test showed no difference between the 2 groups either. The 2 groups differed in the use of aids during walking at 6 weeks, but this did not result in a difference between groups on the gait parameters. It was therefore justified to perform the analyses with the overall group.

DISCUSSION

The purpose of this study was to describe gait recovery after THA based on spatiotemporal gait parameters measured by an ambulatory measuring device while using a test protocol that includes walking at 3 speeds, walking with an additional cognitive task, and a walking endurance test. Our hypothesis was that this test protocol provides measures of gait recovery that cannot be detected when only walking at a preferred speed is assessed.

The results of the study showed that the test protocol provided added value over measurement of only preferred speed. Six weeks after surgery, the results of walking at a preferred speed indicate a very small increase in walking speed and step length and a small decrease in gait variability (variability coefficient). A larger increase in speed and step length is first seen after 6 months. In contrast, already after 6 weeks, the speed–step length relationship during walking at different speeds shows a small but consistent increase in speed and step length (see fig 1). In addition, the assessments of walking at a fast speed show a large speed and step length increase between 6 weeks and 6 months after surgery. In accordance with findings of others,25 our data of walking at different speeds showed that gait variability is speed dependent before surgery (see fig 2). Six weeks after surgery, no decrease of gait variability with speed is seen; instead, gait variability increases slightly. However, after 6 months, an overall decrease of gait variability can be seen and again gait variability decreases with increasing speed. Some caution is needed when interpreting these data because most patients used walking aids 6 weeks after surgery, which might have influenced gait variables.24 All of these findings would not have been obtained if only gait at 1 speed had been analyzed, as is done in most of the previous research.1,4,12-14,25-27 Kyriazis and Rigas26 did analyze gait parameters at 2 speeds; they compared preferred and fast speed of 25 patients preoperatively and 1 and 8 to 10 years after the THA with a conductive walkway system by using telemetry. They concluded that fast walking did not provide information that was not already shown in preferred gait. However, our results show that analyzing gait by assessing walking at 3 different speeds offers more insight in recovery of gait function after THA than assessments of preferred walking alone. As is also shown by our results on gait variability, an additional advantage of assessing gait at different speeds is that when other aspects of gait are studied (eg, muscle activities, angular displacements, trunk accelerations), the test protocol can take into account the speed dependency of these aspects.16,29

Walking while performing an additional attention-demanding cognitive task is a means to measure whether patients are able to walk automatically.4 To our knowledge, this is the first study that uses this measure in THA patients. Analysis is based on the principle of dual-task interference, which is the worsening of performance of the main task (eg, walking) as a result of simultaneously performing an attention-demanding cognitive task (eg, counting backward). Automaticity of the main task is reflected by a low or absent dual-task interference effect. Measuring the interference effect over time gives an indication of the level of returning automaticity.4 The dual-task principle has been used in other orthopedic patient populations (eg, after limb-saving surgery of the lower limb).30 Nonautomaticity can
### Table 2: Gait Parameters and Results of the General Linear Model Analysis

<table>
<thead>
<tr>
<th>Patient Group (N=63)</th>
<th>Preoperatively</th>
<th>6 Weeks</th>
<th>6 Months</th>
<th>GLM RM (main time effect)</th>
<th>GLM RC (preop to 6wk)</th>
<th>GLM RC (6wk to 6mo)</th>
<th>Reference Group (n=19)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walking at slow speed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speed (m/s)</td>
<td>0.74±0.16</td>
<td>0.80±0.17</td>
<td>0.89±0.16</td>
<td>F_{2,124}=37.12 P&lt;.001; partial \eta^2=0.37</td>
<td>F_{1,62}=8.41 P=.005; partial \eta^2=0.12</td>
<td>F_{1,62}=36.96 P=.001; partial \eta^2=0.37</td>
<td>0.92±0.13</td>
</tr>
<tr>
<td>Step length (m)</td>
<td>0.51±0.08</td>
<td>0.55±0.08</td>
<td>0.58±0.09</td>
<td>F_{1,62}=38.03 P&lt;.001; partial \eta^2=0.38</td>
<td>F_{1,62}=25.66 P&lt;.001; partial \eta^2=0.29</td>
<td>F_{1,62}=28.71 P&lt;.001; partial \eta^2=0.20</td>
<td>0.62±0.05</td>
</tr>
<tr>
<td>Step duration (s)</td>
<td>0.71±0.10</td>
<td>0.71±0.09</td>
<td>0.65±0.07</td>
<td>F_{1,62}=16.17 P&lt;.001; partial \eta^2=0.21</td>
<td>F_{1,62}=0.11 P=.74; partial \eta^2=0.00</td>
<td>F_{1,62}=10.35 P&lt;.002; partial \eta^2=0.14</td>
<td>0.68±0.07</td>
</tr>
<tr>
<td>Variability coefficient (%)</td>
<td>13.1±5.8</td>
<td>10.4±4.6</td>
<td>10.5±4.0</td>
<td>F_{1,62}=7.43 P=.001; partial \eta^2=0.11</td>
<td>F_{1,62}=10.35 P&lt;.001; partial \eta^2=0.14</td>
<td>F_{1,62}=4.00 P&lt;.00; partial \eta^2=0.00</td>
<td>8.2±2.7</td>
</tr>
<tr>
<td>Walking at preferred speed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speed (m/s)</td>
<td>0.93±0.19</td>
<td>0.95±0.17</td>
<td>1.14±0.18</td>
<td>F_{1,62}=56.16 P&lt;.001; partial \eta^2=0.48</td>
<td>F_{1,62}=0.99 P=.32; partial \eta^2=0.02</td>
<td>F_{1,62}=105.74 P&lt;.001; partial \eta^2=0.63</td>
<td>1.32±0.15</td>
</tr>
<tr>
<td>Step length (m)</td>
<td>0.56±0.09</td>
<td>0.58±0.08</td>
<td>0.63±0.10</td>
<td>F_{1,62}=40.14 P&lt;.001; partial \eta^2=0.39</td>
<td>F_{1,62}=7.86 P=.007; partial \eta^2=0.11</td>
<td>F_{1,62}=44.59 P&lt;.001; partial \eta^2=0.42</td>
<td>0.71±0.07</td>
</tr>
<tr>
<td>Step duration (s)</td>
<td>0.61±0.07</td>
<td>0.62±0.07</td>
<td>0.55±0.04</td>
<td>F_{1,62}=30.08 P&lt;.001; partial \eta^2=0.33</td>
<td>F_{1,62}=0.38 P=.54; partial \eta^2=0.01</td>
<td>F_{1,62}=69.12 P&lt;.001; partial \eta^2=0.53</td>
<td>0.54±0.03</td>
</tr>
<tr>
<td>Variability coefficient (%)</td>
<td>12.0±5.6</td>
<td>10.5±3.6</td>
<td>9.6±3.1</td>
<td>F_{1,62}=7.55 P=.001; partial \eta^2=0.11</td>
<td>F_{1,62}=6.27 P=.02; partial \eta^2=0.09</td>
<td>F_{1,62}=2.69 P&lt;.11; partial \eta^2=0.04</td>
<td>8.1±3.2</td>
</tr>
<tr>
<td>Walking at fast speed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speed (m/s)</td>
<td>1.14±0.26</td>
<td>1.13±0.21</td>
<td>1.37±0.25</td>
<td>F_{1,62}=57.11 P&lt;.001; partial \eta^2=0.48</td>
<td>F_{1,62}=0.01 P=.73; partial \eta^2=0.00</td>
<td>F_{1,62}=141.43 P&lt;.001; partial \eta^2=0.70</td>
<td>1.73±0.17</td>
</tr>
<tr>
<td>Step length (m)</td>
<td>0.62±0.11</td>
<td>0.63±0.09</td>
<td>0.69±0.11</td>
<td>F_{1,62}=44.71 P&lt;.001; partial \eta^2=0.42</td>
<td>F_{1,62}=1.20 P=.28; partial \eta^2=0.02</td>
<td>F_{1,62}=87.23 P&lt;.001; partial \eta^2=0.59</td>
<td>0.80±0.08</td>
</tr>
<tr>
<td>Step duration (s)</td>
<td>0.55±0.07</td>
<td>0.56±0.06</td>
<td>0.51±0.05</td>
<td>F_{1,62}=32.21 P&lt;.001; partial \eta^2=0.34</td>
<td>F_{1,62}=2.15 P=.15; partial \eta^2=0.03</td>
<td>F_{1,62}=82.02 P&lt;.001; partial \eta^2=0.57</td>
<td>0.47±0.03</td>
</tr>
<tr>
<td>Variability coefficient (%)</td>
<td>10.7±4.8</td>
<td>10.9±4.0</td>
<td>9.0±3.3</td>
<td>F_{1,62}=6.72 P=.002; partial \eta^2=0.10</td>
<td>F_{1,62}=0.21 P=.65; partial \eta^2=0.00</td>
<td>F_{1,62}=14.89 P&lt;.001; partial \eta^2=0.19</td>
<td>11.0±5.9</td>
</tr>
<tr>
<td>Walking with additional task</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speed (m/s)</td>
<td>0.79±0.21</td>
<td>0.85±0.21</td>
<td>0.99±0.21</td>
<td>F_{1,62}=40.62 P&lt;.001; partial \eta^2=0.40</td>
<td>F_{1,62}=6.65 P=.012; partial \eta^2=0.10</td>
<td>F_{1,62}=52.37 P&lt;.001; partial \eta^2=0.46</td>
<td>1.09±0.16</td>
</tr>
<tr>
<td>Step length (m)</td>
<td>0.56±0.09</td>
<td>0.58±0.08</td>
<td>0.62±0.10</td>
<td>F_{1,62}=19.28 P&lt;.001; partial \eta^2=0.24</td>
<td>F_{1,62}=8.20 P=.006; partial \eta^2=0.12</td>
<td>F_{1,62}=20.51 P&lt;.001; partial \eta^2=0.25</td>
<td>0.67±0.07</td>
</tr>
<tr>
<td>Step duration (s)</td>
<td>0.73±0.14</td>
<td>0.71±0.11</td>
<td>0.63±0.07</td>
<td>F_{1,62}=21.94 P&lt;.001; partial \eta^2=0.26</td>
<td>F_{1,62}=1.42 P=.24; partial \eta^2=0.02</td>
<td>F_{1,62}=41.03 P&lt;.001; partial \eta^2=0.40</td>
<td>0.63±0.10</td>
</tr>
<tr>
<td>Variability coefficient (%)</td>
<td>14.0±8.7</td>
<td>11.7±5.9</td>
<td>10.0±4.0</td>
<td>F_{1,62}=7.85 P=.001; partial \eta^2=0.11</td>
<td>F_{1,62}=4.40 P=.04; partial \eta^2=0.07</td>
<td>F_{1,62}=4.93 P&lt;.03; partial \eta^2=0.07</td>
<td>7.4±3.8</td>
</tr>
</tbody>
</table>

**Note.** Values are mean ± SD. Abbreviations: GLM RC, general linear model repeated contrasts; GLM RM, general linear model repeated measures.
have an impact on daily living; gait changes caused by performing an additional cognitive task while walking are associated with increased risk of falling among older adults.

Our results indicate that the inclusion of a dual task is valuable as a means to determine (recovery of) gait automaticity. Particularly the variability coefficients showed improved gait during dual-task performance. Already after 6 weeks, gait variability is reduced. After 6 months, gait variability has further reduced and does no longer differ from the variability that can be expected based on the chosen gait speed (see fig 2). During dual-task conditions, gait speed, but not step length, is reduced in comparison to preferred walking. However, this reduction in gait speed is smaller after 6 months than preoperatively or after 6 weeks. Thus, all our results indicate a return of gait automaticity.

The results of the walking endurance test did not show the expected decrease in speed and step length and increase in gait variability when the first and final 20-m distances were compared. The results also did not show the expected improvement after surgery. Existing differences, although statistically significant, were very small and cannot be considered as clinically relevant. Whether or not a patient completed this test part, the results from the longer walking distance could already be predicted from the results of walking at preferred speed.

Results of the dual-task test part presented an unexpected interference that has not been reported in other similar studies: the association between walking speed and step length deviates from the association seen in the other test parts. The tendency to enunciate a number at each step resulted in a slow gait with relatively large step lengths, which can be because of the rhythmic character of additional task interfering with the walking rhythm. In gait research, there are different attention-demanding tasks that vary in complexity and suitability for different populations and age groups. Gait research on patients recovering from limb-saving surgery of the lower limb has used the auditory Stroop test. Further research is necessary to determine the applicability of this test in THA patients; some consider this task less appropriate for older adults who often have difficulties hearing. Moreover, it is advised that patients also perform the cognitive task as a single task (eg, while sitting down) so that performance on the task with and without walking can be compared too. In our study, we found that performing an additional cognitive task affects gait parameters. In theory, patients can also perform worse on the additional task and not on walking, despite the instructions to direct attention to the counting task while walking. This might be a sound “safety-first” strategy, but with this strategy changes in performance of the cognitive task remain unrecognized.

The walking endurance test appeared not to be a measure of endurance. Either the test has to be prolonged, which makes it difficult to incorporate into daily clinical practice, or the instruction should be given to walk as fast as possible, which is done in the protocol of the 6MWT. We considered this as not representative of independent functioning in daily life. However, comparison with previous studies performed by using the 6MWT is therefore impossible. We do have to keep in mind that the measurements were assessed in relatively healthy patients because only patients without other health problems could participate in the short-stay program. Additionally, 6 patients could not perform the test. These patients did complete the slow-, preferred-, and fast-walking test parts and performed worse on all parameters at all 3 measurement times than the included patients. A remarkable observation during the endurance test was that patients walked faster than during the walking at preferred-speed test part. This occurred despite the fact that participants were instructed to walk at their preferred speed in both test parts. The walking endurance test might be experienced as a competition. Thus, despite instructions, some
participants may want to cover as many meters within the 6 minutes as possible. Another explanation is that walking the 20-m walkway several times back and forth is experienced differently than walking it once. It has been observed that gait speed on a long walkway (eg, 60 m) is higher compared with short walkways.33

The administration of the measurements was easy and well accepted by all patients. None reported any discomfort caused by the portable measuring device so any possible influence thereof on gait can be neglected. The total duration of a measurement was approximately 20 minutes and can therefore be integrated into routine clinical practice.

Replacing the walking endurance test with a “walking on different ground surfaces” test part seems relevant and can be the object of future research.7,16 Patients often complain about problems when walking on uneven pathways, especially after surgery. This can be the result of balance problems because it is known that balance during gait is altered by severe osteoarthritis, and some of these balance problems seem to remain after THA.34 Stability during walking should therefore be assessed while walking under circumstances that challenge the postural control system to obtain an indication of patients’ ability to respond to multiple, irregular perturbations.35 With the ambulant measuring device, tests can be performed outside the laboratory. Hence, gait can be studied under more complex, unpredictable, and real-life circumstances, and this may reveal essential information,7 as is also indicated by the results of our dual-task test part.

CONCLUSIONS

In this study, a test protocol consisting of walking at 3 different speeds while performing an additional cognitive task and an endurance test was used to obtain insight into recovery of gait function after total hip replacement. The hypothesis that the test protocol would provide measures of gait recovery that cannot be detected when only preferred walking speed is assessed was confirmed by a part of our results. First, the use of different speed instructions made it possible to analyze gait parameters, especially the link between walking speed and step length and gait variability over a range of speeds. Already after 6 weeks, the speed–step length relation during walking at different speeds shows a small but consistent increase in step length.42 Differences in performance during walking at different speeds and step length can be a result of balance problems because it can be the result of balance problems because it is known that balance during gait is altered by severe osteoarthritis, and some of these balance problems seem to remain after THA.34 Stability during walking should therefore be assessed while walking under circumstances that challenge the postural control system to obtain an indication of patients’ ability to respond to multiple, irregular perturbations.35 With the ambulant measuring device, tests can be performed outside the laboratory. Hence, gait can be studied under more complex, unpredictable, and real-life circumstances, and this may reveal essential information,7 as is also indicated by the results of our dual-task test part.

Acknowledgments: We thank Sander Heikens and Rienk van der Slikke, from McRoberts BV, for their technical assistance.

References


Suppliers
b. Version 12.0.1; SPSS Inc, 233 S Wacker Dr, 11th Fl, Chicago, IL 60606.
Muscle Tone in Diabetic Polyneuropathy Evaluated by the Quantitative Pendulum Test

Chou-Ching K. Lin, MD, PhD, Ming-Shaung Ju, PhD, Han-Wei Huang, MD


Objective: To quantify the muscle tone of upper limbs in patients with diabetic polyneuropathy (DPN).

Design: Case-control study. Quantitative upper-limb pendulum tests were conducted, and model analysis was performed.

Setting: Outpatient clinic of a medical center.

Participants: The experimental group consisted of patients with type 2 diabetes suffering from symptomatic but not disabling polyneuropathy. The diagnosis of polyneuropathy was based on symptoms, signs, and nerve conduction velocity (NCV) study. The control group consisted of age- and sex-matched normal subjects. In total, 181 subjects were recruited, including 128 controls and 53 DPN patients.

Interventions: Not applicable.

Main Outcome Measures: Quantitative biomechanic parameters (number of swings, relaxation index, stiffness constant and damping coefficient) were formulated and the differences between groups were investigated.

Results: The number of swings and stiffness constant showed no difference between groups. Relaxation index increased and damping coefficient decreased significantly in the DPN group.

Conclusions: Muscle tone, defined as passive resistance in the tested range, in DPN patients was shown to be decreased. The decrease was mainly due to a decrease in the velocity-dependent component. The decrease did not correlate with the decrease in NCV and could not be detected by the conventional manual pendulum test performed at the bedside.

Key Words: Diabetic neuropathies; Muscle hypotonia; Rehabilitation.

© 2007 by the American Congress of Rehabilitation Medicine and the American Academy of Physical Medicine and Rehabilitation

M ANY PATIENTS WITH DIABETIC polyneuropathy (DPN) complain about joint or muscle stiffness. Because polyneuropathy causes hyporeflexia and muscle wasting, however, it is natural to think that DPN leads to hypotonia. Whether the muscle tone is actually increased or decreased is important for management of DPN patients. Unfortunately, there were insufficient data to answer this question.

Muscle tone is the passive resistance of muscle under traction and is the joint display of neural control, muscle status, and connective tissue properties. More attention has been paid to hypertonia, because it is related to many common diseases, such as stroke and Parkinsonism, and can be evaluated semiquantitatively by manual tests. Furthermore, hypertonia can be quantified by electronic pendulum test and mechanical passive stretch.1,2 Hypotonia is less emphasized, because hypotonia can only be inferred from indirect evidence, such as hyperextensibility of joints and pendulum test. There is no recognized standard for quantifying hypotonia and this partly accounts for the relative lack of studies in investigating the muscle tone in DPN patients.

In previous studies,3,4 we designed a simple accessory apparatus to assist performing the pendulum test in the elbow joint and also proposed a biomechanic model of the elbow to formulate parameters for quantification. Initial results indicated that the parameters could differentiate between spasticity in stroke patients and normal muscle tone of healthy subjects.4 In another study, we also established the effects of age, sex, and body weight on muscle tone of normative subjects.5 The main purpose of the present study was, using our developed apparatus and analytic tools, to quantitatively investigate the effects of DPN on muscle tone. The results of this study not only quantify the muscle tone of DPN patients but also have implications for plans of treatment.

METHODS

Participants

This was a case-control study. We recruited subjects from the outpatient clinic of National Cheng Kung University Hospital (NCKUH). For the DPN group, subjects consisted of consecutive referrals to our neurology clinic due to DPN caused by type 2 diabetes from year 2000 to 2005. Diabetes mellitus was defined by the post fasting blood glucose level, glycosylated hemoglobin level, and the continued use of diabetes-specific medications. Polyneuropathy was defined by the symptoms and the nerve conduction velocity (NCV) studies of upper and lower limbs. We deliberately excluded stage 3 patients6 (ie, patients with disabling neuropathy), so the subjects in the experimental group all belonged to stage 2 (ie, symptomatic neuropathy). For the control group, subjects were chosen randomly from our database of normative subjects in the same years by matching the age range and sex. The sources of control subjects were (1) family members of patients who came to our hospital with the patients for electrophysiologic examinations and (2) patients that came for electrophysiologic examinations due to other problems, such as lumbosacral radiculopathy and headache. Subjects with a history of stroke, Parkinsonism, polyneuropathy due to other etiologies, or other neurologic diseases (particularly, those producing abnormal muscle tone, restricted range of motion or decreased muscle power) were excluded. Inclusion criteria for the control group included clear consciousness and cooperativity. No subject was
taking antispasticity medication or muscle relaxant at the time of study. The study protocol was approved by the NCKUH ethics committee on human subject study. Before an experiment, the purpose, the potential hazards and the procedure of the experiment were fully explained to the subjects. A written permission form was signed. The body weight, forearm length, and maximal forearm circumference were measured for the estimation of mass, center of mass, center of gyration, and inertia of the forearm and hand.7,8

Study Design

The experimental setup and procedure was identical to that adopted in a previous study.5 In brief, an accessory apparatus was specifically designed to facilitate performing the pendulum test in the elbow joint (fig 1). The accessory apparatus consisted of a shaft, a weight, and a wrist fastening part. The steel shaft was connected at the mid point to the test bed through a pure rotary joint. An electronic goniometer at this joint measured the elbow joint angle. A weight was fastened to the lower end of the shaft to increase the total inertia and counterbalance the weight of the forearm. The subject lay comfortably on the test bed. The wrist was fixed to the accessory apparatus through the wrist fastening part. The upper part of the shaft was hooked to the test bed with a chain of predesigned length, such that the elbow joint angle was 130° (referencing full extension as 0°). Surface electromyography of the biceps and triceps brachii was collected with 2 pairs of standard cup electrodes and amplified 1000 times after band-pass filtering (1.59–300Hz) using a Polygraph 360 system.a The data collection was started and the chain was released swiftly without informing the subject. The forearm passively swung due to the weight at the lower part of the apparatus. After the swing motion stopped by visual inspection, the data collection was terminated. Six successful trials were collected.

The signals were sampled at 600Hz for 15 to 25 seconds depending on the duration of swing and stored in a personal computer for off-line analyses. The data collection was accomplished with LabView.b The data were preprocessed and if (1) background electromyographic activity was observed and (2) the angle trajectories were inconsistent during the pendulum test, the data were excluded. Because there was no electromyographic activity, the signals were not further processed.

Data Analyses

To quantify the results, we formulated several parameters. Number of swings and relaxation index (RI) were determined from the averaged angle trajectory,9 where the RI was the ratio of maximal swing angle to the final steady-state angle and number of swings was the number of peaks and troughs during the swing (fig 1). In general, both number of swings and RI decreased as the muscle tone increased. Then, the angle trajectory was fitted to a previously proposed biomechanic model of the elbow joint,4

\[ \ddot{\theta} = -\tau - K(\theta - \theta_0) - C\dot{\theta} \]  

(1)

where \( \theta \) is the elbow joint angle, \( \tau \) is the gravitational torque, \( K \) is the stiffness coefficient of tissues, including muscles, around the elbow joint, \( \theta_0 \) is the threshold angle, and \( C \) is the damping coefficient (also known as the coefficient of viscosity). From equation 1, it is clear that \( K \), related to the angle, is a measure of tendency to return to \( \theta_0 \) position and \( C \), related to the angular velocity, is a measure of resistance to swing. In general, \( K \) is independent of muscle tone and is more related to the stiffness of tissue and \( C \) increases as muscle tone increases.4 The parameters (\( K, C \)) were estimated by recursive optimization technique. The goodness of parameter estimation was evaluated with root mean square error (RMSE) between the actual and estimated elbow angle trajectories. The full mathe-
Table 1: Anthropometric Data of Subjects

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Control</th>
<th>DPN</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y)</td>
<td>58.1±11.1</td>
<td>58.4±9.5</td>
<td>.768</td>
</tr>
<tr>
<td>Sex (men/women)</td>
<td>80/48</td>
<td>34/19</td>
<td>.615</td>
</tr>
<tr>
<td>Body weight (kg)</td>
<td>63.9±9.3</td>
<td>67.6±12.1</td>
<td>.016</td>
</tr>
<tr>
<td>Forearm circumference (cm)</td>
<td>26.4±2.5</td>
<td>26.8±3.1</td>
<td>.251</td>
</tr>
<tr>
<td>Forearm length (cm)</td>
<td>30.3±1.9</td>
<td>30.8±1.9</td>
<td>.200</td>
</tr>
<tr>
<td>Estimated weight of FA (kg)</td>
<td>1.4±0.20</td>
<td>1.49±0.27</td>
<td>.016</td>
</tr>
<tr>
<td>Estimated inertia of FA (kg/m²)</td>
<td>0.10±0.02</td>
<td>0.11±0.03</td>
<td>.015</td>
</tr>
</tbody>
</table>

NOTE. Values are mean ± SD or as otherwise indicated. Abbreviation: FA, forearm and hand together.

Statistical Analyses

First, we calculated the group means and standard deviations (SDs) of the 4 quantitative parameters and tested the significance of difference between the 2 groups by Mann-Whitney U test with a significance level of α equal to .05. Next, the relationship among factors and parameters were evaluated and clustered by principal components analysis. Third, K and C as a linear function of weight and DPN were fitted by the multivariate regression technique. We also calculated the significance of correlation between factors by Fisher r to z transformation. Last, we computed the correlation between the parameters and NCV results. StatView was used for the above-mentioned statistical analyses.

RESULTS

One hundred eighty-one subjects (128 controls, 53 DPN patients) completed the pendulum test. The basic data of subjects are listed in table 1. The estimated inertia and total weight of forearm and hand are also listed. Subjects were not matched on body weight and the DPN group was slightly heavier. Motor and sensory NCV was only performed in the DPN group. The NCV was slower than the normative data from our electrophysiology lab (table 2).

An example of the angle trajectory during the pendulum test along with the simulation result is shown in figure 2. A summary and comparison of parameters are listed in table 3. Although there was no significant difference between the 2 groups for parameters number of swings and K, the differences for both RI and C were significant. The pooled mean and SD of RMSE in model fitting was 2.9°±1.5°.

We did principal components analysis to evaluate the relationship among parameters and the respective contribution of the clustered factors (table 4). The results show that the parameters were clustered into 4 factors. Factor 1 (consisting mainly of number of swings, RI, and C) contributed 30.8%, factor 2 (consisting mainly of weight, forearm length, and sex) 32.4%, factor 3 (consisting mainly of age and K) 19.5%, and factor 4 (consisting mainly of DPN) 15.9% to the total variance, respectively. In other words, there were 4 independent clusters of parameters, the first one being related to viscosity, the second one to body weight and sex, the third one to stiffness and age, and the last to polyneuropathy.

According to one of our previous studies, K and C were relatively independent of age, and the effects of sex and forearm length on K and C were related to the difference in body weight. When all the subjects in the 2 groups were pooled together, K was fitted as a function of weight (W) and P (DPN) by stepwise multiple linear regression,

\[ K = 0.039W + 0.182, \]

where P equals 1 represents the control group and P equals 2 represents the DPN group, and \( r^2 \) equals .19 (P<.001). P was dropped because of nonsignificant contribution (P=.28). Similarly, C was fitted as

\[ C = 0.009W - 0.218P + 0.424, \]

where \( r^2 \) equals .09 (P<.001). The effect of interaction between weight and P was checked by adding W×P in the regression analyses and the results showed that the interaction was nonsignificant.

The linear correlations between weight versus K and C (fig 3) were calculated (K: \( r^2=.436, P<.001; \) C: \( r^2=.184, P=.13 \)). The difference between the 2 correlation coefficients was significant (P<.01). The results indicated that K was more dependent on body weight. The linear correlations between parameters and NCV results were very low (fig 4), indicating the correlation between the joint stiffness and NCV results was very weak in the DPN patients that we recruited.

Table 2: NCV Results of DPN Group

<table>
<thead>
<tr>
<th>Parameters</th>
<th>NCV (m/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median motor</td>
<td>47.4±5.3</td>
</tr>
<tr>
<td>Ulnar motor</td>
<td>47.7±6.8</td>
</tr>
<tr>
<td>Median sensory</td>
<td>23.2±17.7</td>
</tr>
<tr>
<td>Ulnar sensory</td>
<td>24.8±18.7</td>
</tr>
</tbody>
</table>

NOTE. Values are mean ± SD.

Table 3: Parameters and Comparisons of 2 Groups

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Control</th>
<th>DPN</th>
<th>z</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of swings</td>
<td>4.10±1.92</td>
<td>4.25±1.71</td>
<td>-0.686</td>
<td>.492</td>
</tr>
<tr>
<td>RI</td>
<td>1.43±0.17</td>
<td>1.56±0.25</td>
<td>-3.257</td>
<td>.001</td>
</tr>
<tr>
<td>K (Nm·s)</td>
<td>2.64±0.88</td>
<td>2.93±0.99</td>
<td>-1.657</td>
<td>.098</td>
</tr>
<tr>
<td>C (Nm·s/rad)</td>
<td>0.78±0.44</td>
<td>0.59±0.29</td>
<td>-2.410</td>
<td>.016</td>
</tr>
</tbody>
</table>

NOTE. Values are mean ± SD.
DISCUSSION

Comparison With Other Studies

As described above, there is very little objective data on muscle tone in DPN. Nielsen et al\textsuperscript{10} reported that ankle joint stiffness decreased in long-term type 1 diabetic patients. The decrease was very subtle. Stiffness in the report was not defined as the elastic component, but as the gross resistance to passive stretch. Nielsen did not investigate the changes in individual stiffness components. Our results also indicated that the total resistance to passive stretch was decreased and, furthermore, the decrease was in the damping component (C in table 3), whereas the elastic component showed no difference from the control group. Duquette et al\textsuperscript{11} studied the viscoelastic properties of knee ligaments in congenital diabetes mellitus rats and reported that the storage compliance (corresponding to 1/K) did not differ significantly from that in the control group but the loss compliance (corresponding to 1/C) was increased. The results were compatible with ours. The authors also commented on many previous studies, noting problems in study design that resulted in the inconsistency of results.\textsuperscript{12,13} Athanasiou et al\textsuperscript{14} studied the effects of diabetes on biomechanic properties of human ankle cartilage and found that the cartilage became softer. The authors did not analyze the changes in stiffness components, either.

Many studies\textsuperscript{15,16} reported that the range of motion (ROM) of joints decreased in diabetic patients, which seemed to be contradictory to the trend of hypotonia and our results. This point is discussed in detail below.

Factors Contributing to Muscle Tone

Neural control, muscle status, and biomechanic properties of connective tissue are 3 main factors affecting the magnitude of muscle tone. Traditionally, neural control has been thought to be the most dominant factor. Polyneuropathy affects mainly the spinal segmental reflexes. Stretch reflexes with afferents from muscle spindles and Golgi tendon organs are thought to play a key role in the spinal segmental neural control of muscle tone. The effects of muscle status on muscle tone are less well defined. Though it is known that a larger muscle produces a larger muscle force, no data show whether, in the passive state, C is larger in a larger muscle or joint. Wiegner and Watts,\textsuperscript{17} in a small series, showed that K of the elbow joint correlated linearly with the volume of the arm. Chleboun et al\textsuperscript{18} also reached a similar conclusion by measuring elbow flexor volume and angular stiffness of the elbow joint. We calculated the correlation between weight versus K and C, and the results indicated that K was more dependent on body weight. When

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
<th>Factor 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forearm length</td>
<td>-.033</td>
<td>.878</td>
<td>-.150</td>
<td>.053</td>
</tr>
<tr>
<td>Body weight</td>
<td>.108</td>
<td>.800</td>
<td>.125</td>
<td>.130</td>
</tr>
<tr>
<td>Sex</td>
<td>.082</td>
<td>-.762</td>
<td>.107</td>
<td>.130</td>
</tr>
<tr>
<td>Age</td>
<td>-.269</td>
<td>-.397</td>
<td>.726</td>
<td>.118</td>
</tr>
<tr>
<td>No. of swings</td>
<td>.991</td>
<td>-.043</td>
<td>.247</td>
<td>-.259</td>
</tr>
<tr>
<td>Ri</td>
<td>.689</td>
<td>.001</td>
<td>-.089</td>
<td>.320</td>
</tr>
<tr>
<td>C</td>
<td>-.731</td>
<td>.001</td>
<td>.310</td>
<td>-.151</td>
</tr>
<tr>
<td>K</td>
<td>.289</td>
<td>.400</td>
<td>.804</td>
<td>-.006</td>
</tr>
<tr>
<td>DPN</td>
<td>&lt;.001</td>
<td>.066</td>
<td>.086</td>
<td>.947</td>
</tr>
</tbody>
</table>

NOTE. Bold face denotes when the absolute value of the relative weighting is greater than .06.


Fig 4. Relationship between motor NCV of median nerve and K and C, respectively, in the DPN group.
the product of forearm length and circumference was used as an indicator of forearm volume, the correlation of this indicator versus K and C was similar to the above conclusion. Last, though the resistance to the passive stretch is defined as the muscle tone, in addition to muscle, connective tissue around the joint is also stretched in the testing process. More and more data indicate that the biomechanic properties of muscle and connective tissues around the joint, including tendons and ligaments, also affect the grossly perceived muscle tone. The extreme case is the complete joint contracture. One study showed that muscle tone and joint ROM were related, that is, smaller ROM was associated with hypertonia.

In the case of DPN in this study, spinal segmental reflexes, including stretch reflexes, decreased, contributing to decreased muscle tone. Diabetes, through biochemical actions, changes the molecular interactions of connective tissues and also contributes to the decreased muscle tone. The relative immobilization, however, due to poor physical condition may cause contracture of joints, leading to a smaller ROM, and increased passive resistance. We did not match body weight in the control and study groups. The greater body weight of the DPN groups may also have contributed to elevated tone. The paradox between decreased ROM and decreased muscle tone may have been secondary to the movement range over which the test was performed. In measuring ROM, whole ROM was evaluated, whereas in our pendulum test, the pendular swing was from 130° to about 60°, that is, part of the extension was not tested. In summary, many factors with variable degree of influence contributed to the final gross manifestation of muscle tone. In the DPN subjects in our study that we chose and the movement range over which the joint was tested, the factors leading to decreased muscle tone dominated over those that may have led to increased tone.

As described in the Methods section, all subjects in the DPN group were rated as having stage 2 diabetic polyneuropathy. We expect that muscle tone in some DPN patients may change from hypotonia to hypertonia as polyneuropathy becomes more severe and the joints are more immobilized and develop contracture.

Comments on the Parameter Results

Although the parameters number of swings, RI, and C were clustered together as a factor in the principal components analysis, number of swings, in contrast to RI and C, did not show a significant difference between the 2 groups. The possible explanation is that number of swings, representing the number of swings and being an integer, was a discrete variable, whereas RI and C were continuous variables. Discrete variables were less powerful in detecting subtle changes. Because number of swings was not increased in the DPN group, it indicates that the conventional manual pendulum test performed at the bedside is ineffective in detecting the subtle hypotonia in DPN. Le Cavorzin et al showed, by model simulation, that the weight of forearm and hand has an effect on angle trajectory similar to C. In other words, RI can be increased by increasing the weight of forearm and hand without changing C. Because weight of forearm and hand was estimated from body weight by multiplying a constant, we tested the effects of body weight on the RI by similar model simulations (fig 5). First, the pendulum angle trajectory of the control group was calculated by incorporating the mean K, C, weight, and forearm length of the control group. Then, weight was changed by ±10% and ±20% and simulations were repeated. For ±20% change in W, the changes in the RI are only 3.7% (1.42±0.03). If we increased weight by 20% and assumed the variance of the RI was identical, the statistical difference between the RI of the control and DPN groups was still significant. Because the difference in the mean body weight between the 2 groups was only 6%, the above results indicated that the difference between the RI of the control and DPN groups was not due to the difference in body weight. Parameter RI has the advantage of being independent of model and easy to calculate. On the other hand, C has a definite physical meaning and interpretation.

Clinical Implications

One of our previous studies showed that C was increased and K remained constant in stroke patients with spasticity. This study showed that C was decreased and K remained constant in DNP patients. Combined, these results strengthen the concept that testing of muscle tone is more stretch-velocity–dependent and that the damping coefficient, C, can be a quantitative indicator of muscle tone. In contrast, the stiffness constant, K, intuitively representing the stiffness of the joint, is not a primary factor in determining muscle tone. In the previous study, it was also shown that C was linearly correlated with the scores on the Modified Ashworth Scale. Because we have found no comparable scale or other relevant quantitative studies about hypotonia in the past, we do not know the linear correlation between the clinical severity of hypotonia and C. To our knowledge, this is the first quantitative study of hypotonia.

In addition to the application of these findings to DPN patients, accurate muscle tone quantification can also clarify the effects of many drugs on muscle power and muscle tone. The decomposition of damping (velocity-dependent) and stiffness (position-dependent) components by model analysis assists in the search of underlying mechanisms of altered muscle tone in pathologic conditions.

CONCLUSIONS

The present study showed that muscle tone, defined as the passive resistance in the tested range, was decreased in stage 2
DPN patients. The decrease was mainly due to a decrease in the velocity-dependent component.

References

Suppliers
a. NEC, 7-1, Shiba 5-chome, Minato-ku, Tokyo 108-8001, Japan.
c. SAS Institute Inc, 100 SAS Campus Dr, Cary, NC 27513.
A Comparison of Psychometric Properties of the Smart Balance Master System and the Postural Assessment Scale for Stroke in People Who Have Had Mild Stroke

Chi-Wen Chien, BS, Ming-Hsia Hu, PhD, Pei-Fang Tang, PhD, Ching-Fan Sheu, PhD, Ching-Lin Hsieh, PhD


Objective: To compare the psychometric properties (including the test-retest reliability, responsiveness, and predictive validity) of the Smart Balance Master (SBM) system and the Postural Assessment Scale for Stroke patients (PASS) in patients with mild stroke.

Design: One repeated-measures design (at a 2-wk interval) was used to examine the test-retest reliability of the SBM and PASS, and another similar design was applied to investigate their responsiveness. Patients who participated in the responsiveness study were followed up approximately 1 year later, and the predictive validity of the SBM system and PASS were examined by assessing the patients’ comprehensive activities of daily living (ADL) function.

Setting: Three rehabilitation units in Taiwan.

Participants: Twenty patients with chronic stroke in the reliability study; 40 and 32 patients who had recently had a stroke in the responsiveness and predictive validity studies, respectively.

Interventions: Not applicable.

Main Outcome Measures: Three computerized tests of the SBM (the equilibrium score of the Sensory Organization Test, scores in rhythmic weight-shifting tests, and scores in the limits of stability test) and the PASS were used. The combination of the Barthel Index and Frenchay Activities Index was used to represent the comprehensive ADL function.

Results: For the SBM, all but the weight-shifting tests of the SBM had moderate to high reliability (intraclass correlation coefficient [ICC] range, .78–.91). The responsiveness of the equilibrium score and the limits of stability test were moderate (effect size [d], .63) and small (d range, .27–.33), respectively, whereas the responsiveness of the weight-shifting tests were limited (d range, .04–.29). All but the weight-shifting tests of the SBM in the second evaluation had acceptable predictive validity for comprehensive ADL function (r² range, .15–.17). The PASS showed high reliability (ICC=.84) and small responsiveness (d=.41), and the PASS in the second evaluation had acceptable predictive validity (r²=.24).

Conclusions: The PASS and the equilibrium score and limits of stability scores of the SBM had acceptable test-retest reliability, responsiveness, and predictive validity in patients with mild stroke, but the psychometric properties of the weight-shifting tests of the SBM should be further examined before consideration of their usage in patients with stroke.

Key Words: Cerebrovascular disorders; Posture; Psychometrics; Rehabilitation.

© 2007 by the American Congress of Rehabilitation Medicine and the American Academy of Physical Medicine and Rehabilitation

A patient’s balance ability after stroke is an important component toward attaining autonomy in activities of daily living (ADLs).1–3 Measuring balance ability in patients with stroke assists clinicians in the assessment of the severity of the stroke and selection of the most appropriate treatment. The use of balance measures generally relies on either laboratory instrumentation or clinical scales, each of which has its own clinical strengths. Laboratory instrumentation such as the Smart Balance Master (SBM) system,4 allows clinicians to quantify standing balance ability with high resolution, whereas clinical balance scales have practical advantages that include ease of administration and low cost. To be clinically useful, however, all balance measuring tools must be scientifically sound in terms of 3 basic psychometric properties: reliability, validity, and responsiveness.5–7

Test-retest reliability is a fundamental requirement of balance measures.7,8 Balance measures with acceptable test-retest reliability allow users to determine whether the changes in patients’ balance abilities are the result of specific treatment or of measurement inconsistency. Responsiveness or sensitivity to change is another requirement for balance measures designed to evaluate change.9,10 If a measure is unable to detect change in balance ability, it will fail to indicate the extent of change in patients who have had improvement in balance function. Last, balance measures with good predictive validity are essential to assessing the predictive outcome of functional recovery.1 The importance of predicting patients’ functional outcome is that it enables clinicians to address many issues such as selecting appropriate treatment programs, setting treatment goals, and facilitating a suitable discharge plan.11,12 Thus, to ensure that all balance measures are clinically useful, it is critical to examine their test-retest reliability, predictive validity, and responsiveness.

With regard to the psychometric properties of the SBM, the limits of stability test of the SBM has shown high test-retest reliability in patients with mild stroke.12 Fair to good test-retest reliability for the Sensory Organization Test (SOT) of the SBM was previously established in community-dwelling elderly...
The study protocols were approved by the ethics committee of the participating hospitals.

**Procedure**

To investigate the test-retest reliability of the SBM and the PASS, all subjects with chronic stroke were evaluated initially using the SBM and the PASS and were then reassessed after a 2-week interval. On the SBM evaluations, each subject was allowed to have 1 practice to learn how to perform the tests of the SBM. One therapist conducted all the SBM measurements and another therapist administered the PASS. The SBM and the PASS were administered (in a random order) within a 24-hour period.

To investigate the responsiveness of the SBM and the PASS, we assessed all of the subjects with recent stroke by using the SBM and the PASS twice at a 2-week interval. The administration procedures of the SBM and the PASS in each assessment were the same. During the 2 weeks in the responsiveness study, each subject received intensive rehabilitation programs in the hospital 5 days a week. The programs included at least 2 hours of intervention each day, which routinely provided the subjects with therapeutic tasks involving postural control training and weight shifting. About 1 year after the first evaluation, the Barthel Index and Frenchay Activities Index (FAI) were administered to these same patients or his/her main caregiver in the responsiveness study via telephone interview by the third therapist. The third therapist was not informed of the assessments performed by the previous 2 therapists. The follow-up test administration was completed in order to determine the predictive values of the SBM and PASS on the first 2 evaluations for comprehensive ADL function (combining the Barthel Index and FAI) about 1 year after the first evaluation.

**Instrument**

**Balance measures.** The PASS was specifically developed to measure balance function in patients with stroke, including those with very poor performance. The PASS contains 12 four-point items that gauge a person’s balance performance in situations of varying difficulty, namely, maintaining or changing a sitting, lying, or standing position. Its total score ranges from 0 to 36. The psychometric properties of the PASS were found to be satisfactory in patients with stroke during the first 6 months after stroke.

The SBM has been used to quantify standing balance abilities in patients with stroke. The components of the SBM are dual forceplates, an overhead bar (safety straps, carabiners), and associated equipment. The SBM provides objective measures of the basic balance components involved in balance control, including the center of gravity, postural alignment, postural stability, limits of stability, and rhythmic weight shift. Three commonly used computerized tests of the SBM were used in this study, including the SOT, the limits of stability, and the rhythmic weight-shifting tests.

The SOT is a method commonly used to quantify stance stability under varying sensory conditions. The SOT is designed to measure a subject’s ability to maintain equilibrium, while systematically altering the input information available to the somatosensory and/or visual systems. In this study, the SOT consisted of 20-second trials under 6 sensory conditions: (1) eyes open and fixed support, (2) eyes closed and fixed support, (3) sway-referenced vision and fixed support, (4) eyes open and sway-referenced support, (5) eyes closed and sway-referenced support, and (6) sway-referenced vision and support. The equilibrium score is a quotient calculated from the actual peak-to-peak sway angle and the theoretical maximum
sway angle to indicate postural stability. The equilibrium score used in this study was calculated based on the average of the following 14 scores: 1 score for each of conditions 1 and 2, and 3 scores for each of conditions 3 to 6. Higher equilibrium scores indicate better balance.

The limits of stability test involved shifting the weight to 8 target positions arranged in an ellipse on the monitor screen, the perimeter of which corresponded to 50% of the theoretical limits of stability. The default of the perimeter in the limits of stability test is usually set at 75% of the limits of stability in the SBM, which appears to be difficult for patients with stroke according to our clinical observation. The 75% of the limits of stability was thus reduced to 50% in this study. During the limits of stability test, subjects were required to shift their center of gravity to follow a ball-shaped cursor to each target as it was highlighted, and to remain at that target position for 4 seconds before returning to the center of the ellipse. Targets were highlighted in order, and each target was selected only once. The time allowed to reach a target was 8 seconds. The average movement time (limits of stability time; in seconds) and path sway (limits of stability path; in percentage of limits of stability) were determined across targets. A score of 100% in the limits of stability path indicated that no path deviation occurred (ie, perfect execution), whereas scores other than 100% indicated the presence of path deviation. The weight-shifting tests involved rhythmically shifting weight side to side to 50% of the limits of stability, calculated based on an inverted pendulum model from the subject’s body height, at 3-second and then at 2-second pacings. Subjects were instructed to match the timing and movement of a ball on the screen by shifting their body weight side to side to the target lines denoted on the screen in the mediolateral rhythmic weight shift test. The forward-backward weight shift test was also performed. Data reflecting the average magnitude of the movement path (expressed as a percentage of the limits of stability) were produced from 6 trials at each pace, for each movement direction. The absolute error relative to the targets (set at 50% of the limits of stability) was calculated by subtracting 50 from the score obtained (50% indicating perfect execution) and recording it in absolute terms. Lower absolute errors indicated better abilities in rhythmic weight shifting. Further details on the SBM can be found elsewhere.

**Comprehensive ADL measure.** We used the Barthel Index (BI) to evaluate 10 basic ADL items, with the total score ranging from 0 to 100. The Barthel Index has been shown to be a reliable, valid, and responsive measure of disability, and the FAI was developed as a means of measuring social activities or lifestyle after stroke. The FAI includes 15 items that are associated with normal social activities, and these items were each rated from 0 to 3 points. The range of possible scores is 0 to 45. The FAI has been shown to be a reliable and valid measure of instrumental ADLs in patients with stroke.

Previous studies found that the Barthel Index and FAI scores could be combined to represent a comprehensive ADL function, representing the entire continuum of disability. Hsueh et al. proposed that combining all but 2 items (social occasions, walking outside) of the Barthel Index and FAI, and simplifying the multiple response categories into dichotomous categories, can create a Rasch-transformed score as a comprehensive ADL function. The Rasch-transformed logit score of the combination of the Barthel Index and FAI, which can be considered interval-level measurement, was used in this study.

**Other clinical measures.** We also administered the MMSE to determine the clinical characteristics of patients with stroke in the study. The MMSE was developed to assess cognitive status. The MMSE was used to ensure that the patients in this study had no serious cognitive impairment. A cutoff score of 20 was chosen to adjust for the possible impacts of age and level of education in this study.

**Data Analysis**

To facilitate comparison between the psychometric properties of the PASS and those of the SBM, we followed previous studies and used parametric statistical procedures to analyze the PASS scores.

**Test-retest reliability.** We used 3 statistical indices to investigate the test-retest reliability of the balance measures over the 2-week period. First, paired t tests were performed to examine the changes for statistical significance. Second, a 1-way random effects model intraclass correlation coefficient (ICC) was used to summarize the strength of the test-retest reliability. ICC values 0.8 or higher indicate high reliability, and ICC values in the range of 0.6 to 0.8 represent moderate reliability. Third, the minimal detectable change (MDC) at both the individual level and the group level was established. The MDC is the threshold value that determines whether a subject’s (or a group of subject’s) score changes in the balance measures of the responsiveness study were beyond chance variation in measurement (ie, denoting true change). The MDC with a confidence level of 95% (MDC95) at the individual level (MDC95,ind) was calculated by the following formula:

\[ \text{MDC95,ind} = 1.96 \times \text{the standard deviation of the baseline score} \times \sqrt{2(1 - \text{ICC value of the test = retest reliability})} \]

The MDC at the group level (MDC95,group) was determined by the MDC95,ind value divided by the square root of the size of the sample. Changes smaller than the MDC95,ind or MDC95,group cannot be reliably interpreted as “true changes” in the score for a single subject or the mean score for a group, respectively.

**Responsiveness.** To examine the changes in scores of the balance measures over the 2-week period, we calculated the effect size (d) by dividing the mean change scores by the standard deviation of the baseline scores in the same subjects. Cohen suggested that an effect size greater than 0.8 is large, 0.5 to 0.8 is moderate, and 0.2 to 0.5 is small.

**Predictive validity.** We determined the predictive validity of the balance measures by the strength of the association between the first 2 evaluations of the SBM and PASS, and the comprehensive ADL function assessed at 1 year after the first evaluation. The association was examined by using the squared Pearson correlation coefficient (r²). The r² value between 2 measures indicates the extent of the explanatory (predictive) power. Greater than 10% predictive power (ie, r²>0.1) with statistical significance (P<.05) was considered as the minimal threshold to determine the predictive validity of the balance measures.

**RESULTS**

In the test-retest reliability study, a total of 20 patients with chronic stroke were included. Table 1 shows demographic and clinical characteristics of these subjects in the study. An independent sample of 42 patients who had experienced a recent stroke was recruited in the responsiveness study. Two patients declined to return for the second evaluation, because one lived too far away from the hospital and the other felt dizzy during the first evaluation of the SBM and thus was reluctant to be reassessed; therefore, 40 patients participated in the responsiveness study. Of the remaining 40 patients with stroke, 32 patients participated in the follow-up evaluation of the predic-
The ICC for the PASS over the 2-week interval was .84 (95% confidence interval, .64 to .93). The responsiveness of the weight-shifting tests was highly variable (ICC range, .24 to .72), whereas the responsiveness of the weight-shifting tests was low or not responsive (d range, .04 to .29). The PASS had small responsiveness (d = .41). These results indicated that the PASS and the equilibrium score and limits of stability tests of the SBM were more responsive than the weight-shifting tests of the SBM.

In addition, the subjects’ change scores among the balance measures showed a significant difference between the measurements (paired t test, P > .05).

### Responsiveness

Table 3 compares the responsiveness among the balance measures. For the SBM, the responsiveness in all tests of the SBM was somewhat inconsistent. The equilibrium score was moderately responsive (d = .63) and the limits of stability tests had small responsiveness (d range, .27 to .33), whereas the responsiveness of the weight-shifting tests was low or not responsive (d range, .04 to .29). The PASS had small responsiveness (d = .41). These results indicated that the PASS and the equilibrium score and limits of stability test of the SBM were more responsive than the weight-shifting tests of the SBM.

Table 2: Comparison of the Test-Retest Reliability and the MDC Among Balance Measures Used in This Study

<table>
<thead>
<tr>
<th>Measures</th>
<th>First Evaluation (mean ± SD)</th>
<th>Second Evaluation (mean ± SD)</th>
<th>Mean Difference Score</th>
<th>ICC* (95% CI)</th>
<th>MDC95,ind</th>
<th>MDC95,group</th>
<th>P*</th>
</tr>
</thead>
<tbody>
<tr>
<td>PASS</td>
<td>32.8 ± 2.0</td>
<td>32.9 ± 1.9</td>
<td>0.15</td>
<td>.84 (.64 to .93)</td>
<td>2.22</td>
<td>0.50</td>
<td>.56</td>
</tr>
<tr>
<td>SBM</td>
<td>61.5 ± 11.9</td>
<td>65.2 ± 12.0</td>
<td>3.65</td>
<td>.81 (.54 to .93)</td>
<td>14.39</td>
<td>3.22</td>
<td>.03</td>
</tr>
<tr>
<td></td>
<td>4.8 ± 2.1</td>
<td>4.6 ± 2.0</td>
<td>-2.21</td>
<td>.91 (.79 to .96)</td>
<td>1.75</td>
<td>0.39</td>
<td>.30</td>
</tr>
<tr>
<td></td>
<td>454.0 ± 211.4</td>
<td>443.9 ± 211.8</td>
<td>-0.86</td>
<td>.78 (.53 to .91)</td>
<td>274.86</td>
<td>61.46</td>
<td>.75</td>
</tr>
<tr>
<td></td>
<td>22.8 ± 10.6</td>
<td>22.4 ± 9.3</td>
<td>-0.33</td>
<td>.71 (.36 to .88)</td>
<td>15.83</td>
<td>3.73</td>
<td>.86</td>
</tr>
<tr>
<td></td>
<td>19.2 ± 7.1</td>
<td>18.1 ± 11.0</td>
<td>-1.03</td>
<td>.51 (.01 to .80)</td>
<td>13.78</td>
<td>3.56</td>
<td>.67</td>
</tr>
<tr>
<td></td>
<td>13.7 ± 9.2</td>
<td>11.4 ± 8.0</td>
<td>-0.21</td>
<td>.24 (.26 to .64)</td>
<td>22.23</td>
<td>5.39</td>
<td>.40</td>
</tr>
<tr>
<td></td>
<td>10.9 ± 8.4</td>
<td>10.9 ± 8.4</td>
<td>-0.10</td>
<td>.38 (.23 to .72)</td>
<td>18.32</td>
<td>4.44</td>
<td>.70</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; FB, forward and backward weight shift; LR, left to right weight shift; 3-s, 3-second pacing; 2-s, 2-second pacing.

*One-way random effects model ICC.

*MDC at the individual level = \(1.96 \times \text{standard deviation of the baseline score} \times \sqrt{\frac{1}{2} \left[1 - \text{coefficient of test-retest reliability}\right]}\).

*MDC at the group level = \(\text{MDC}_{95,\text{ind}} / \sqrt{\text{size of sample}}\).

*Paired t tests.
level, 15% of the subjects had change scores higher than the MDC95,ind value on the PASS. For the SBM, 30% and 10% of the subjects had change scores higher than the corresponding MDC95,ind on the equilibrium score and limits of stability tests, respectively, but varied from none to 12% in the weight-shifting tests. Moreover, the mean change scores (see table 3) of the PASS and all but the weight-shifting tests of the SBM at the group level were much smaller than their corresponding MDC95,group values (see table 2).

Predictive Validity

Table 4 shows the association between the first 2 evaluations of the SBM and the PASS within 3 months after stroke and the comprehensive ADL function at 1 year after the first evaluation. Only the limits of stability time \( r^2 = .14 \) and the mediolateral weight-shifting test at 3-second pacings \( r^2 = .16 \) in the first evaluation were significantly associated with the comprehensive ADL function. All tests of the SBM in the second evaluation, except for the weight-shifting tests, had acceptable explanatory power to predict the comprehensive ADL function \( (r^2 \text{ range}, .15 - .18) \). The explanatory power of the PASS to predict the comprehensive ADL function was also acceptable \( (r^2 = .24) \) in the second evaluation.

**DISCUSSION**

The psychometric properties of laboratory balance instrumentation are commonly thought to be better than those of clinical balance scales, although direct comparisons between both types of balance measures have been lacking. To the best of our knowledge, this study is the first to compare the test-retest reliability, responsiveness, and predictive validity of the SBM, which is a laboratory balance instrument, with those of the PASS, which is a clinical balance scale, on the same group of patients with stroke who could stand independently. The comparison was conducted to thoroughly investigate the psychometric properties of the SBM and PASS and to provide evidence for clarifying the issue on the comparison of the psychometric properties of balance instrumentation and clinical balance scales. This evidence will be helpful in assisting clinicians and researchers to select an appropriate balance measure based on the stringency of the psychometric properties, rather than the type of the balance measure.

**Test-Retest Reliability**

Our results showed that the equilibrium score, the limits of stability test (ie, the limits of stability time, limits of stability path), and the PASS had moderate to high reliability. Furthermore, the PASS was found to have superior test-retest reliability to the tests of the SBM in the study. Liston and Brouwer\(^2\) also found that the test-retest of the BBS was not inferior to that of the SBM in patients with stroke. These observations indicate that balance instrumentation may not have test-retest reliability superior to that of clinical balance scales. It is also noted that the mean difference score of the equilibrium score over a 2-week interval was significantly higher, and thus learning effects of the equilibrium score seemed to exist between both measurements.

The poor test-retest reliability of the weight-shifting tests at 3- and 2-second pacings was found in our study as well as in 2 previous studies,\(^2,43\) indicating that the weight-shifting tests were unreliable balance measures. The reason why the weight-shifting tests had poor test-retest reliability might be that these tests were too difficult for these patients to achieve consistent performances. For example, 3 subjects in the first and 2 in the second evaluation showed difficulties to shift their body weight to match a ball moving at 3- and 2-second pacings; therefore, the SBM did not score their performances. Perhaps more practice given before the weight-shifting tests could enable patients with stroke to familiarize themselves with weight shifting, thus resulting in improved test-retest reliability.

**Responsiveness**

The responsiveness of the SBM has rarely been examined. We found that the equilibrium score and limits of stability of the SBM and also the PASS showed acceptable responsiveness, and that the equilibrium score revealed higher responsiveness than the PASS. Possible learning effects of the equilibrium score seemed to exist between both measurements.

**Table 3: Comparison of the Responsiveness Among Balance Measures Used in This Study**

<table>
<thead>
<tr>
<th>Measure</th>
<th>First Evaluation (mean ± SD)</th>
<th>Second Evaluation (mean ± SD)</th>
<th>Mean Change Score</th>
<th>*Effect size.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PASS (n=40)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equilibrium score (n=40)</td>
<td>51.0±15.6</td>
<td>60.9±14.2</td>
<td>9.83 .63</td>
<td></td>
</tr>
<tr>
<td>Limits of stability time (n=40)</td>
<td>5.0±1.3</td>
<td>4.6±1.4</td>
<td>−0.36 .27</td>
<td></td>
</tr>
<tr>
<td>Limits of stability path (n=40)</td>
<td>521.7±248.2</td>
<td>439.6±180.5</td>
<td>−82.04 .33</td>
<td></td>
</tr>
<tr>
<td>Weight-shifting FB 3-s (n=40)</td>
<td>25.6±9.0</td>
<td>26.5±7.0</td>
<td>0.93 .10</td>
<td></td>
</tr>
<tr>
<td>Weight-shifting FB 2-s (n=33)</td>
<td>24.9±9.1</td>
<td>22.1±7.7</td>
<td>−2.69 .29</td>
<td></td>
</tr>
<tr>
<td>Weight-shifting LR 3-s (n=37)</td>
<td>15.4±8.1</td>
<td>13.7±9.1</td>
<td>−1.87 .25</td>
<td></td>
</tr>
<tr>
<td>Weight-shifting LR 2-s (n=32)</td>
<td>11.8±9.2</td>
<td>11.4±7.7</td>
<td>−0.35 .04</td>
<td></td>
</tr>
</tbody>
</table>

*Effect size.

**Table 4: Relationships Between the SBM and PASS at Both Evaluations and Comprehensive ADL Function at 1 Year After the First Evaluation**

<table>
<thead>
<tr>
<th>Measure</th>
<th>First Evaluation Predicting Comprehensive ADL Function* ( r^2 )</th>
<th>Second Evaluation Predicting Comprehensive ADL Function ( r^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PASS (n=32)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equilibrium score (n=32)</td>
<td>.07</td>
<td>.17†</td>
</tr>
<tr>
<td>Limits of stability time (n=32)</td>
<td>.14†</td>
<td>.18†</td>
</tr>
<tr>
<td>Limits of stability path (n=32)</td>
<td>.04</td>
<td>.15†</td>
</tr>
<tr>
<td>Weight-shifting FB 3-s (n=32)</td>
<td>.01</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Weight-shifting FB 2-s (n=28)</td>
<td>.01</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Weight-shifting LR 3-s (n=30)</td>
<td>.16†</td>
<td>.05</td>
</tr>
<tr>
<td>Weight-shifting LR 2-s (n=28)</td>
<td>.09</td>
<td>.04</td>
</tr>
</tbody>
</table>

*As measured by the combined Barthel Index and FAI.
†Significant at \( P<.05 \).
reliability part of the study. The learning effects might have led to inflation in the responsiveness of the equilibrium score. In addition, the limits of stability test was found to have higher responsiveness than the weight-shifting tests, but the responsiveness of both limits of stability and weight-shifting tests was inferior to that of the equilibrium score as well as the PASS. These observations indicate that the responsiveness of the PASS is comparable to that of the SBM.

In addition, the MDC value of a measure can be used as a threshold to determine whether observed changes are true (ie, beyond measurement error). We found that only a few patients (≤12% of the sample) had a change score over the MDC group values of the weight-shifting tests. The mean change scores of the PASS and all but the weight-shifting tests of the SBM were much smaller than their corresponding MDC group values. These observations indicate that it is difficult for the patients to reach a true change (beyond the measurement error) in the weight-shifting tests, and that the weight-shifting tests have the smallest responsiveness among the balance measures. The 2-week interval in the design of the responsiveness study might have been too short, however, in that only a small balance change in the patients could be found in this study. Future studies that adopt a longer time frame (eg, 6wk) to ensure a larger change may be needed to further compare the responsiveness of the balance measures.

Predictive Validity

The balance abilities of patients soon after a stroke have been found to be closely associated with long-term functional improvement. Few studies have examined the predictive validity of the SBM and compared it with that of the clinical balance scales in patients with stroke. We found that the PASS and the tests of the SBM, with the exception of the weight-shifting tests, at an early stage (ie, within 3mo) could substantially predict comprehensive ADL function at 1 year after first assessment \((r^2>0.1)\). In particular, the PASS showed a predictive validity comparable to those offered by the equilibrium score and the limits of stability test of the SBM used in the study. The predictive power of the PASS and the SBM for comprehensive ADL function found in this study, however, may not have clinical utility due to their low predictive powers \((r^2≤.24)\). The low predictive validity for ADL function in the PASS and SBM might be expected, because ADL function of a patient depends on multiple factors (eg, balance, cognition, motor function, age, multiple factors). Therefore, the low predictive validity in the present study might indicate that the SBM and the PASS could be used complementarily to measure balance abilities for patients with stroke. The psychometric properties of the weight-shifting tests of the SBM still require thorough examination before further use.

Study Limitations

The present study has 3 limitations. First, the modest sample size of this study, and particularly some missing data in the weight-shifting tests, might threaten the generalization of this study. Future studies that recruit more subjects may be needed to further validate our findings. Second, we excluded patients who could not follow instructions. It was determined that patients with severe cognitive dysfunction might not perform well on complicated balance instrumentation (ie, the SBM). The psychometric properties of the SBM are not suitable for generalization to patients with cognitive dysfunction. Third, this study only focused on the comparison of the psychometric properties between the SBM and PASS in measuring balance. In addition to measuring balance, the SBM possesses other valuable functions, such as analyzing the neuromuscular mechanism with respect to postural strategies or providing training protocols. For example, the weight-shifting tests were found to have poor psychometric properties in this study, but the weight-shifting tests might be used as balance training tools to strengthen subjects’ rhythmic, reciprocal movement in order to meet the timing demands in daily activities. Therefore, further studies to compare other functions of the SBM with clinical balance scales are needed to explore the clinical practicability in stroke rehabilitation.

CONCLUSIONS

The results of this study showed that the PASS and the equilibrium score and the limits of stability test of the SBM had acceptable test-retest reliability, responsiveness, and predictive validity in patients with stroke without cognitive dysfunction or visual impairment who could stand independently. The psychometric properties of the weight-shifting tests of the SBM should be further explored before considering their usage in patients with stroke.

References


Objectives: To investigate acute changes in the biceps tendon after a high-intensity wheelchair propulsion activity and to determine whether these changes are related to subject characteristics.

Design: The biceps tendon was imaged with ultrasound before and after wheelchair basketball or quad rugby. The average diameter of the tendon was calculated as well as the echogenicity ratio (the pixel intensity ratio of the biceps tendon to a reference just superficial to the tendon sheath).


Participants: Forty-two subjects who participated in wheelchair basketball or quad rugby at the Veterans Games.

Interventions: Not applicable.

Main Outcome Measures: Biceps tendon diameter and biceps echogenicity.

Results: The echogenicity ratio of the tendon significantly decreased from 1.97 to 1.73 after the event (P = .003). The diameter of the biceps tendon increased from 4.60 to 4.82mm (P = .004). Also, it was found that the change in tendon diameter positively correlated with the time of play (P = .004).

Conclusions: Acute changes in biceps tendon properties after exercise were found and likely represent edema, a first sign of overuse injury. The significance of continuous activity was shown by the fact that subjects who had more playing time showed a larger increase in tendon diameter.

Key Words: Rehabilitation; Shoulder; Tendon injuries; Ultrasonography; Wheelchairs.

© 2007 by the American Congress of Rehabilitation Medicine and the American Academy of Physical Medicine and Rehabilitation

ORIGINAL ARTICLE

Ultrasound Imaging of Acute Biceps Tendon Changes After Wheelchair Sports

Stefan van Drongelen, PhD, Michael L. Boninger, MD, Bradley G. Impink, Tagreed Khalaf, MD

Among people with a spinal cord injury (SCI), shoulder pain is a common complaint with a prevalence of between 30% and 73%. Injuries to the shoulder can be the result of an acute overload, multiple-repetition load (chronic overuse), or single-event trauma to tissue made vulnerable by overuse (acute on chronic). Persons who depend on a manual wheelchair for their main means of mobility experience these high loads and repetitive overuse loads on their upper extremities in locomotion and in performing daily tasks.

To diagnose shoulder pathology, different imaging methods have been used including magnetic resonance imaging, magnetic resonance arthrography, and computed tomography. Musculoskeletal ultrasound has the advantages over these systems of being easy to perform in an office setting and allowing dynamic examination of joints during motion. Ultrasound has been performed for approximately 25 years, and the technique has been refined and is widely accepted as an important tool with high precision and accuracy in diagnosis. A study by Allen and Wilson has shown ultrasound to be a very powerful and accurate method for examining the biceps mechanism, rotator cuff, and synovium.

Read and Perko and Thain and Adler have shown that ultrasound is an excellent clinical tool to diagnose chronic tendon injuries. Acute tendon injuries have not been investigated. These acute changes to tendons of the shoulder, which might occur after sport participation, can be part of the pathologic process that leads to chronic pathology and pain. Acute overuse injuries to tendons normally begin with inflammation of the tissue surrounding the tendon, after which fluid is secreted into the tendon sheath. Isolated biceps tendinitis is uncommon. Biceps tendinitis is often secondary to impingement or rotator cuff tears because the biceps tendon sheath is in direct communication with the glenohumeral joint. Insight into the acute changes in the biceps tendon after activity could contribute to a better understanding of the etiology of chronic pathology.

When using ultrasound, structures containing more fluid appear darker or are less echogenic. Abnormal or nonuniform echogenicity on ultrasound is more often indicative of tendinosis, tendinopathy, or mucoid degeneration.

There is already evidence that links the load of daily wheelchair propulsion to chronic overuse symptoms. During sport participation, the load on the upper extremities is more pronounced because the forces to propel the wheelchair and the frequency of these forces are much higher. Next to the frequency and the magnitude of the forces, the time of play and the number of rest periods are relevant factors to the development of musculoskeletal disorders. Not only are the task requirements risk factors for shoulder damage but personal characteristics, like body mass index, are risk factors as well. The rolling resistance, and thus the work needed to propel a wheelchair, is directly related to weight. Besides, in wheelchair basketball and quad rugby, there is a lot of starting and stopping, resulting in an increased frequency of accelerating. When accelerating, body mass is a critical component for the necessary force.
In this study, we used ultrasound to investigate acute changes in the biceps tendon after a high-intensity wheelchair propulsion activity. We hypothesized that because of increased edema, the biceps tendon would become less echogenic and increase in size after participation in an intense physical activity. Furthermore, we expected that subjects with a higher body mass would show larger changes in the biceps tendon characteristics and that the duration of wheelchair activity is positively correlated to these changes.

METHODS

Participants

Persons who participated in wheelchair basketball or quad rugby at the National Veterans Wheelchair Games of 2004 and 2005 were recruited for this study. A convenience sample of 42 subjects participated in this study after giving written informed consent.

Subjects were eligible to participate if they used a wheelchair as their main means of mobility and were between 18 and 65 years old. The exclusion criterion was a history of trauma or surgery to both arms. The protocol used to examine the structures of the shoulder was based on previously described techniques.\(^20-23\) The protocol of this study received approval by the institutional review board of the Veterans Affairs, Pittsburgh Healthcare System.

Procedure

All participants completed a general information form that provided date of birth, date of injury, injury level, sex, height, and weight (table 1), and answered a questionnaire related to their shoulder and arm pain. Subjects then underwent a baseline ultrasound of the nondominant arm before the event and an ultrasound within 30 minutes after the event. The nondominant side was measured to limit the influence of daily tasks, which are more likely performed by the dominant arm.

All shoulders were examined by 1 of 2 examiners who had prior training and experience with musculoskeletal ultrasound. Ultrasound was performed with a Diasus Ultrasound Scanning Systems\(^4\) using a 5- to 10-MHz linear array transducer.

The protocol used to examine the structures of the shoulder was based on previously described techniques.\(^20-23\) The transverse image of the biceps tendon was obtained when the subject’s arm was resting in his/her lap. Supination of the hand with external rotation of the shoulder improved the visualization of the bicipital groove. The transducer was then turned 90° to obtain the longitudinal image of the biceps tendon.

At the postgame ultrasound, participants reported their actual time of play in the event. In addition, time from the end of the event until the postgame ultrasound was noted.

Data Analysis

Before analyzing the images, the images were screened for usability by 2 reviewers. Base criteria were good definition of the boundaries of the biceps tendon and identification of the same part of the tendon in both images. In case of doubt, a third reviewer was consulted. To prevent erroneous values being included in the results, the calculated values were checked for normal distribution by probability plots. Values deviating more than 4 standard deviations (SDs) were considered outliers and were removed from the analyses.

The biceps tendon characteristics (tendon diameter, echogenicity) were analyzed with a Matlab\(^b\) program, which was written by the investigators. The images were displayed in a random order, providing for blinding of the pre- and postgame images.

In the longitudinal images of the tendon, a length of 2cm was selected by the investigator (fig 1), and the tendon was outlined over this length. This selection included the narrowest part of the tendon. The average diameter of this selection was calculated.

To investigate changes in the fluid content of the tendon, the echogenicity of the tendon was determined by calculating the average pixel intensity value (a pixel intensity value of 0 corresponded to black and 255 to white) throughout the selected area. It was possible that we observed a decrease in tendon echogenicity when comparing the 2 images, but this could be inaccurate without considering all information in the entire image. If the overall brightness (controlled by machine settings) of the first image was greater than that of the second image, we would see a decrease in echogenicity not related to a change in the tendon. To resolve this issue, we used the tendon to reference echogenicity ratio. By using the average pixel intensity of a reference area just superficial to the tendon (one tenth of the total space above the tendon) (see fig 1), we controlled for differences in the overall brightness as well as differences caused by probe orientation or signal variation with depth. The echogenicity ratio was calculated as the pixel intensity ratio of the biceps tendon to a reference just superficial to the tendon sheath. A lower echogenicity ratio means that the tendon is darker with respect to the reference and, therefore, contains more fluid/water than the reference. If this ratio decreases, then there are 3 possible explanations: (1) the tendon became darker (ie, more fluid in the tendon); (2) the reference became lighter (ie, less fluid in the reference); or (3) both 1 and 2 (ie, the tendon became darker and the reference became lighter). The opposite would be true if the ratio increased.

Statistical Analysis

Paired sample t tests were performed to find differences between the biceps tendon diameter and the tendon echogenicity ratio of the pre- and postgame ultrasound images. A 1-way analysis of variance (ANOVA) was performed to find differences between the images of the persons with tetraplegia, those with paraplegia, and those without an SCI (independent variable: disability with 3 levels: tetraplegia, paraplegia, non-SCI; dependent variables: biceps tendon properties, subject charac-

<table>
<thead>
<tr>
<th>Subjects with tetraplegia (n=11)</th>
<th>Age (y)</th>
<th>Height (m)</th>
<th>Body Mass (kg)</th>
<th>Injury Level (range)</th>
<th>Time Since Injury (y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>45.8±9.1</td>
<td>1.81±0.09</td>
<td>76.7±10.3</td>
<td>C1 incomplete to C7 complete</td>
<td>20.2±9.6</td>
<td></td>
</tr>
</tbody>
</table>

NOTE. Values are mean ± SD. Abbreviation: NA, not applicable. *Significantly different between non-SCI and subjects with paraplegia (P=.038).
teristics). A Levene test of homogeneity of variance was performed to validate the application of ANOVA.

An independent t test was performed to find differences in the biceps tendon properties between subjects reporting pain or no pain in the measured shoulder during the previous month. Correlations between biceps tendon properties, subject characteristics, and play time were evaluated by using a Spearman ρ.

Significance was set at P less than .05 and a trend was reported with a P less than .10.

RESULTS

Participants

Of the 42 subjects who entered the study, data of 34 subjects (33 men, 1 woman) were used for the data analysis (see table 1). The images of 5 subjects were unsatisfactory to clearly identify and measure the biceps tendon. One subject did not participate in his sporting event; therefore, no postgame ultrasound could be performed. The biceps characteristics of 2 subjects were more than 4 SDs from the mean; therefore, these subjects were considered outliers and were omitted as well. We compared excluded subjects to those included and found no difference with respect to age, years of injury, or height. However, the excluded subjects had a significantly higher body mass.

The Levene test of homogeneity of variance showed that it was justified to compare the groups with an ANOVA. The subjects without an SCI had a significantly higher body mass compared with the subjects with paraplegia (F=3.625, P=.038). Subjects with tetraplegia trended toward more years since injury compared with non-SCI subjects (F=2.973, P=.066).

The average time from the end of the event to undergoing the postgame ultrasound was 10.1±6.9 minutes. The average time the subjects reported participating in their sporting event was 28.7±16.6 minutes.

Biceps Tendon Characteristics

The echogenicity of the tendon in relation to the reference superficial to the tendon sheath significantly decreased from 1.97 to 1.73 after the event (t=2.160, P=.038). The diameter of the biceps tendon increased from 4.60 to 4.82 mm, but this change was not significant (t=−1.377, P=.178).

Also, for the biceps tendon characteristics, the Levene test of homogeneity of variance showed that it was justified to use an ANOVA. No differences were found for the echogenicity ratio and the tendon diameter before and after the event among the 3 subject groups (table 2).

Four subjects with tetraplegia and 9 subjects with paraplegia reported pain during the last month in the measured arm. Subjects who reported pain showed a lower echogenicity ratio before (t=2.480, P=.019) as well as after the event (t=2.056, P=.048) compared with subjects who did not report pain.

Correlations

A Kolmogorov-Smirnov test showed that play time and time after play were nonuniformly distributed so a nonparametric correlation test was used. The only correlation found between the biceps tendon properties and the subject characteristics was that the echogenicity (tendon/reference) ratio of the pregame images was negatively correlated with weight (ρ=−.363, P=.035). The biceps diameter trended toward being larger in subjects with a higher body mass (ρ=.298, P=.087).

Playing time trended toward being related to change in diameter (ρ=−.303, P=.082), with greater playing time positively related to change in biceps diameter. On inspection of the data, it was noted that all subjects who reported participa-

Table 2: Biceps Tendon Characteristics Before and After the Event

<table>
<thead>
<tr>
<th>Subjects</th>
<th>Diameter Before</th>
<th>Diameter After</th>
<th>Echogenicity Before</th>
<th>Echogenicity After</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjects with tetraplegia (n=11)</td>
<td>4.19±0.61</td>
<td>4.42±0.79</td>
<td>2.01±0.68</td>
<td>1.80±0.56</td>
</tr>
<tr>
<td>Subjects with paraplegia (n=21)</td>
<td>4.77±0.99</td>
<td>4.98±0.93</td>
<td>1.95±0.82</td>
<td>1.66±0.56</td>
</tr>
<tr>
<td>Non-SCI subjects (n=2)</td>
<td>4.97±1.32</td>
<td>5.27±0.75</td>
<td>1.96±0.16</td>
<td>2.02±0.69</td>
</tr>
<tr>
<td>Mean (n=34)</td>
<td>4.60±0.92</td>
<td>4.82±0.90</td>
<td>1.97±0.74</td>
<td>1.73±0.56</td>
</tr>
</tbody>
</table>

NOTE. Values are mean ± SD.
ing for over 30 minutes (n=8) had an increase in tendon diameter. This was further evaluated with a Kruskall-Wallis test comparing the over 30-minute group to the group that played 30 or less minutes, and a significant difference was found \( (P=.004) \) (fig 2).

**DISCUSSION**

In this study, we found that after exercise the tendon to reference echogenicity ratio decreased. The decrease in echogenicity likely represents an increase in fluid in the tendon. Increased fluid in the tendon can be representative of edema, and one might expect it to be the cause of an increase in tendon diameter. However, the increase in biceps tendon diameter we found (from 4.60 to 4.82mm after the game) was not significant.

In the acute phase of tendon pathology, increased fluid content leads to swelling of the tendon.\(^{25}\) In the chronic phase, small amounts of fluid within the biceps tendon sheath are abnormal but are considered nonspecific. Because the tendon of the long head of the biceps is intra-articular and in direct contact with the glenohumeral joint, larger amounts of fluid can be an indication for abnormalities elsewhere in the glenohumeral joint.\(^{20,26}\) In our subject population, we found that subjects who reported pain showed a lower echogenicity ratio of the biceps tendon, which might indicate that pathology in the shoulder is present.\(^{27-29}\)

The echogenicity ratio of the tendon showed a negative correlation with body mass in the pregame images. Surprisingly, the relation between the echogenicity ratio and body mass was not found in the postgame images. Also, we did not find a relation between the biceps diameter and body mass. Bigger changes in the postgame images were expected for subjects with a higher body mass because the amount of work needed to propel a wheelchair is directly related to weight.\(^{18}\)

With respect to the subject’s participation in the event, we found that the change in tendon diameter positively correlated with the time of play. Thus, in subjects who played longer, the tendon diameter change was larger (ie, the tendon became wider); indicating that duration is a risk factor for developing overuse injuries.

It has already been stated from the beginning of wheelchair studies that this task is responsible for the large amount of overuse injuries in the shoulder.\(^{3,15,30,31}\) However, hard evidence of linking wheelchair propulsion to physical evidence has always been a limitation. In this study, we found not only that the changes in the biceps tendon characteristics are directly related to wheelchair propulsion but also that the amount of change was related to the duration of propulsion. Although the sports activities (quad rugby or wheelchair basketball) might be considered too extreme compared with daily activities, they are probably not when the obstacles and barriers encountered daily in the community (eg, constant acceleration and braking, curbs, slopes) are taken into account.

This study does not show that a short bout of exercise is related to chronic pathology. A recent study by Mercer et al\(^ {14}\) reported that long-term wheelchair use leads to damage. They found that specific joint forces and moments during wheelchair propulsion are directly related to shoulder pathology. Reducing subjects’ wheelchair activities is not preferable because exercise can reduce the risk on coronary heart diseases.\(^ {32}\) So instead of limiting subjects’ participation, the need is to reduce the overall force to propel the wheelchair. This could be accomplished by an alternative wheelchair setup, propulsion training, and weight control.

This study measured wheelchair users at an athletic event, which made it hard to control everything during the pretest and posttest measurements. Therefore, we were extra careful to exclude possible errors in the analyses. First, the images were screened by 2 examiners to identify the boundaries and the same part of the tendon in both images. After the initial analyses, unrealistic values deviating more than 4 SDs were excluded. Excluding these subjects did not lead to more significant results. However, the fact that the excluded subjects had a higher body mass could have biased the result. The signal properties of ultrasound vary with the type of tissue through which the signal is traveling, and the signal quality decreases when traveling through more soft tissue. The difference found in body mass between the excluded subjects and the subjects with a spinal cord injury likely reflect difficulty imaging subjects with a higher body mass.

**Study Limitations**

A limitation of this study was the variability in the personal characteristics of the included subjects. Subject characteristics could not be controlled because the measurements took place at an athletic event. Our main outcome measure was the difference in tendon characteristics before and after the event. This repeated-measure design should control for subject variability. The fact that we had a heterogeneous group adds to the generalizability of our findings.

Another limitation of this study was that not all subjects were measured at the same time after the exercise. The ultrasound was taken as soon as possible after the games. However, several subjects had to be measured and subjects had to come to the provisional laboratory, so the time until they were measured was as long as 30 minutes. Because time after competition did not correlate to the measured ultrasound variables, we believe it is safe to say that measuring within 30 minutes after the game did not influence the results; however, measurements at fixed time intervals after the participation in the sporting event would be recommended for future studies.

To standardize the protocol used even more, it would be ideal to measure subjects in a laboratory not only immediately after the exercise but also to follow subjects during the subsequent hours as well. Because ultrasound is sensitive to the orientation of the probe in relation to the anatomic structure of interest (anisotropy), procedures to guarantee the same orientation of the probe before and after the exercise bout will improve the results.

**CONCLUSIONS**

Acute changes in biceps tendon properties after exercise were found. A decrease in tendon echogenicity likely repre-
sents edema, a first sign of overuse injury. Subjects who had more playing time showed a larger increase in the tendon diameter, showing the significance of continuous activity. The fact that subjects with pain had a lower tendon echogenicity might indicate additional pathology. Further research will be necessary to improve the understanding of the effect of acute changes on chronic shoulder pathology.

References

Suppliers
a. Dynamic Imaging Limited, 9 Cochrane Sq, Brucefield Industrial Pk, Livingston, EH54 9DR, Scotland.
b. The MathWorks Inc, 3 Apple Hill Dr, Natick, MA 01760-2098.
Forced Use as a Potential Cause of Gastrocnemius Tears During Neurologic Rehabilitation: A Report of 2 Cases

Steve R. Fisher, MPT, Laura L. Wiggs, PT, Cindy B. Ivanhoe, MD


Broadly defined, forced-use therapy uses specific techniques designed to engage the patient with brain injury in activities that disallow overcompensation with the noninvolved or less involved body segments while forcing the use of the more involved segments. Some applications may involve placing the patient’s hemiparetic extremity in a closed-chain weight-bearing activity with therapist support. We describe 2 cases of gastrocnemius muscle tears that occurred during inpatient neurologic rehabilitation that may be attributed to forced use of the hemiplegic lower extremity. Each presented with signs and symptoms indicative of deep vein thrombosis of the calf but was later confirmed with magnetic resonance imaging to be muscle tears. Some closed-chain, forced-use activities may be ill advised in the early stages of rehabilitation or if force generation of the muscle is inadequate to provide a protective response to over-stretching. Gastroc-soleus tears should also be considered in the differential diagnosis of unilateral or even bilateral lower-extremity swelling and pain in neurologically impaired patients who are undergoing forced-use therapy.

Key Words: Case report; Leg; Muscles; Rehabilitation; Stroke.

© 2007 by the American Congress of Rehabilitation Medicine and the American Academy of Physical Medicine and Rehabilitation

Forced use as a potential cause of gastrocnemius tears during neurologic rehabilitation...
The patient presented for an outpatient PT evaluation the same day after discharge from the day hospital program. At this time, he complained of right calf pain, which his mother stated had worsened over the past couple of days. The physical therapist observed tenderness to palpation, increase in temperature of the right popliteal fossa and upper calf, and mild swelling of the midcalf. The therapist advised notification of the physician. Considering the patient’s history, he was taken to the local emergency department for workup of possible DVT. A Doppler study was negative, but the emergency department physician did not totally rule out DVT. The patient was non-weight bearing at this time and did not participate in PT. A second Doppler study 3 days later was also negative. Her physician cleared her for weight bearing as tolerated 4 days after the last Doppler study.

The patient continued to have pain in the calf and popliteal fossa, which was worse with weight bearing. A right anterior drawer test was equivocal, and the therapist recommended referral to an orthopedic surgeon for possible anterior cruciate ligament (ACL) injury and gastrosoleus tear. (The patient had a history of right ACL repair 2 years before but no ongoing symptoms related to the right knee before the AVM.) At this time, the patient stated that she had felt a pull in her calf while performing forced use weight-bearing activities while an inpatient just before discharge. The orthopedic surgeon ordered an MRI with contrast of the knee and calf, which revealed intact but surgically repaired ACL and 3 ovoid areas of increased intensity in the gastrocnemius. This was interpreted as gastrocnemius muscle tears. The patient was cleared by the orthopedic surgeon for weight bearing as tolerated and for pain-free stretching of the gastrocnemius. By this point, she had lost range in her ankle and was lacking 25° of dorsiflexion passively. The patient was unable to tolerate weight bearing for gait as her primary means of mobility until approximately 1.5 months after the most likely time of injury.

DISCUSSION

In both cases, the patients were young, highly motivated people engaged in intensive neurologic rehabilitation who participated in similar forced-use activities of treatment. They each developed increased muscle tone in plantarflexor muscle groups and had motor point blocks with phenol. One patient underwent a brief course of serial casting after the motor point blocks. Although we cannot definitively identify the exact time and cause of the gastrocnemius muscle tears in these 2 cases, considering the time course of events, it is highly likely they occurred during neurologic rehabilitation just before their pre-
sentation at our outpatient clinic. It is our belief that they occurred through the application of commonly used, specific forced-use exercises. It is also possible that some combination of forced use, the condition of post neurolytic injections, and serial casting to increase dorsiflexion range may increase susceptibility to muscle-strain injuries in neurologic patients. One patient specifically recalled the forced-use exercise that she believes caused the injury. The other patient had significant aphasia and could not communicate a specific event. He performed the same type of weight-bearing activities with a different therapist immediately before discharge from the rehabilitation day hospital, however. His mother reported that nothing done at home likely contributed to the muscle tear. The patient wore an articulated AFO during all walking outside of therapy as well. His suggested prior history of exogenous anabolic steroid use could have played a role in predisposing the muscle to tearing during heavy use. However, this was never verified and even if so would have been discontinued months earlier.

Reports of muscle tears in the neurologic rehabilitation population are very rare. Although muscle tears typically occur as a result from direct trauma, contusion, or indirect stretch injury, strain thresholds exist for both passive and active injury and most often take place as the result of excessive stretch while the muscle is being activated. Circumstances would have occurred during forced-use activities involving full weight bearing on the forefront of the involved lower limb while standing at the edge of a step or rung of a ladder. Normally, the hindfoot would not be allowed to be unsupported when the involved extremity is in a raised position. Still, it is not clear whether the muscle tears occurred as a result of passive hyper-dorsiflexion or through muscle activation during weight bearing in a non–hyper-stretched position. Although muscle tears are not associated with neurolytic procedures, motor point blocks with phenol have been shown to cause muscle necrosis in animal models. When correlated with the clinical presentations and the likelihood that the radiologists interpreting the studies were aware of the patients’ neurologic and motor point injection history, it is highly unlikely that the findings here represent pure necrosis without tear. It is possible that the reduced muscular activity that occurs after phenol injections diminishes the already impaired natural defense mechanism against overstretching, a decreased gain of the stretch reflex. The gastrocnemius is also particularly susceptible to injury even with intact protective stretch reflexes. It crosses multiple joints and has complex architecture, both conditions known to increase susceptibility to strain injury.

It is widely accepted that recovery after stroke is in part dependent on treatment intensity, with more pronounced effects with higher treatment intensity. Forced-use methods are recognized as valuable treatment tools for therapists working with individuals with a brain injury. Closed-chain, weight-bearing techniques, including those used in these cases, are taught in schools and continuing education courses to physical and occupational therapists as treatment alternatives for appropriate neurologically involved patients. Yet, the literature is notably sparse regarding the acute or long-term effects on the joints and soft tissues of patients encouraged to weight bear, via any imposed method, through a neurologically weakened extremity. In the efforts to provide intense levels of stimulation using forced-use therapy, perhaps a minimal criterion of voluntary muscle activity should be considered before engaging in some closed-chain, weight-bearing exercise. To our knowledge, these minimal criteria have not been established or put forth in the rehabilitation literature. For the lower extremity, we recommend that the patient with hemiparesis be able to perform a supported heel raise on the affected extremity before engaging in forced-use activities that may allow forefoot weight bearing over an unsupported heel or even moderate dorsiflexion. Future research should be directed toward establishing appropriate criteria and guidelines for the safe application of forced-use therapies that involve closed-chain, weight-bearing exercise. There is also a need for investigations into the acute and long-term effects of weight bearing through a neurologically weakened extremity on the joints and soft tissues of that extremity.

CONCLUSIONS

Gastroc-soleus tears should be considered in the differential diagnosis of unilateral or even bilateral lower-extremity swellings in patients undergoing neurologic rehabilitation. Forced use exercises for the hemiparetic lower extremity may place weakened muscles, especially those that cross multiple joints, at increased risk for strain injuries. It is possible that some combination of commonly utilized forced use exercises, motor point blocks, and serial casting contributed to the muscle tears seen in these 2 cases. In regard to the lower extremity, we recommend that the patient with hemiparesis be able to perform a supported heel raise on the affected extremity before engaging in forced use activities that may allow forefoot weight bearing over an unsupported heel or dorsiflexed ankle.

References

The Complications of Scar Formation Associated With Intrathecal Pump Placement

Marina G. Protopapas, DO, Elizabeth Bundock, MD, Susan Westmoreland, VMD, Christopher Nero, MD, W. Andrew Graham, PhD, Shanker Nesathurai, MD


A 40-year-old man had an intrathecal morphine-baclofen pump inserted for the treatment of severe dystonia affecting all limbs and severe low back pain. The etiology of his dystonic symptoms, despite thorough investigations, was uncertain. At age 45, the patient fell resulting in a cervical spinal cord injury. He underwent C2 through C5 instrumentation and fusion for cervical spine stabilization. Subsequently, an intrathecal morphine-baclofen pump was implanted to control pain and decreased spasticity. The patient ultimately died at age 48 from complications of pneumonia, and an autopsy was performed. Gross pathologic examination revealed that the intrathecal catheter entered the posterior aspect of the lumbar thecal sac, but coursed superiorly in the anterior intradural space. The catheter tip exited the thecal sac in the upper thoracic spine and became embedded in a fibrotic scar. Displacement of the catheter tip of the intrathecal morphine-baclofen pump and subsequent formation of scar tissue resulted in decreased drug delivery, contributing to diminished pain control and functional status. Catheter displacement and epidural scar formation must be considered as a potential cause of ineffective pain control and decreased functional status in patients with intrathecal morphine-baclofen pumps.

Key Words: Case report; Intrathecal injections; Muscle spasticity; Spinal injections; Rehabilitation.

© 2007 by the American Congress of Rehabilitation Medicine and the American Academy of Physical Medicine and Rehabilitation

INTRATHECAL DRUG DELIVERY systems are effective in the management of spasticity and intractable pain. Typically, these systems consist of a pump that contains a reservoir of medications such as morphine, clonidine, and baclofen. A catheter is attached to the pump, and the terminal portion is placed in the intrathecal space, thereby delivering medications on a continuous basis. Clinicians, however, should be aware that granulomas can form at the distal end of the catheter tip that decrease drug delivery and/or irritate the exiting nerve roots and/or spinal cord. This complication can result in pro-
at the thoracic level may irritate the exiting nerve roots (ie, intercostal neuralgia) or even mimic cholecystitis by producing right upper quadrant pain. Catheter tip masses in the lumbar region may irritate exiting nerve roots, and thereby mimic such conditions as a herniated lumbar disk and spinal stenosis. Scar formation at the catheter tip may result in insidious compression of the spinal cord or cauda equina and cause slowly progressive weakness. Other early signs of scar formation development include an unexplained increase in muscle spasticity.

The etiology and pathophysiologic mechanisms that lead to the formation of catheter tip inflammatory masses remains unclear. Opioids can incite an inflammatory response in brain and spinal cord tissue through yet to be determined mechanisms. Alternatively, morphine can trigger a mitogen-activated protein kinase cascade, thus activating lymphocyte activity. Opioid compounds may cause human endothelial cells, granulocytes, and monocytes to release nitric oxide, which in turn might, in the presence of mesangial cells, lead to monocyte migration. Whether the phenomenon is \( \mu \)-opioid-receptor mediated or naloxone reversible remains to be determined. A widely accepted strategy to prevent granuloma formation is to maintain the drug dose and concentration as low as possible for as long as possible while still achieving symptom control.

**CONCLUSIONS**

Intrathecal drug delivery systems are potent tools in the management of spasticity and pain. Nevertheless, every treatment has potential side effects and complications. In this context, clinicians should maintain a high index of suspicion to identify granuloma formation as early as possible.

References

**CLINICAL NOTE**

**Musculocutaneous Nerve Injury After Simulated Freefall in a Vertical Wind-Tunnel: A Case Report**

Kenneth Mautner, MD, John C. Keel, MD


We report a case of a skydiver with isolated musculocutaneous nerve injury, which occurred after prolonged positioning of the arm during simulated freefall in a vertical wind-tunnel. Musculocutaneous nerve injury is rare, and the mechanism of isolated injury to this nerve is not entirely understood. Isolated peripheral nerve injuries such as this easily mimic other injuries and can be difficult to diagnose. The skydiver complained of right arm weakness and numbness that began after training in a vertical wind-tunnel. Exam revealed weakness in right elbow flexion and forearm supination, and diminished sensation in the right lateral forearm. Electrodiagnostic testing revealed a decreased amplitude in the right lateral anterbrachial cutaneous nerve sensory nerve action potential, and fibrillations and positive sharp waves in the biceps and brachialis muscles. By 5 months, the subject reported complete sensory and motor recovery. Physical and electrodiagnostic findings corresponded to the distribution of the musculocutaneous nerve. The mechanism of injury was likely the prolonged abducted, extended, and externally rotated position of the shoulder during simulated freefall. Although isolated nerve injuries are uncommon, unusual activities and physiologic demands of athletes can result in such injuries. It is important to be aware of peripheral nerve injuries to facilitate proper diagnosis and management.

**Key Words:** Athletic injuries; Case report; Electrodiagnosis; Musculocutaneous nerve; Rehabilitation.

© 2007 by the American Congress of Rehabilitation Medicine and the American Academy of Physical Medicine and Rehabilitation

**EXTREME SPORTS CAN RESULT** in unusual, even unexpected, injuries, as illustrated by the following case of a recreational skydiver who presented to the sports medicine clinic with an uncommon peripheral nerve injury. In general, it is important for sports medicine practitioners to be aware of peripheral nerve injuries in all athletes, because nerves and supporting structures are at risk due to high physiologic demands. These injuries are more common in the upper limbs and tend to be sport-specific, related to biomechanics. Peripheral nerve injury may not be obvious, however, and may be misdiagnosed or improperly managed.

In the sports medicine literature, musculocutaneous nerve injury is rare, and is usually described in association with other nerve injuries, such as brachial plexus injury. Isolated musculocutaneous nerve injury due to any cause is very rare, and its mechanism is not completely understood. We report an isolated, unilateral musculocutaneous nerve palsy in a recreational skydiver after freefall simulation in a vertical wind-tunnel.

**CASE DESCRIPTION**

A 24-year-old right-handed woman presented to the sports medicine clinic with right arm weakness and numbness. Symptoms began several days before, immediately after she performed repeated sessions of recreational skydiving simulation in a wind-tunnel. Her arms had been in an abducted, extended and externally rotated position for repeated, prolonged durations while she maintained the freefall position. After 1 session, she noticed profound weakness in her right arm, and tingling in the right forearm. She reported no pain in the neck, arm, or shoulder.

Physical exam at the initial visit revealed a flaccid right biceps, with elbow flexion strength grade of 3 and forearm supination strength grade of 4 on the Medical Research Council scale. Manual muscle testing of right forearm pronation, wrist extension and flexion, and hand intrinsics were grade 5. Light touch and sharp sensation were diminished in the right lateral forearm only. Left upper-limb motor and sensory examinations were normative. Muscle stretch reflexes were normative and symmetrical, and there were no long-tract signs. Neck motion and foraminial compression testing were normative. There was no Tinel sign elicited at the elbow or wrist. There was no muscle atrophy, and the head of the humerus was at a normative level. There was no swelling, ecchymosis, or tenderness about the right biceps origins or insertions, and the proximal and distal biceps attachments were intact. Passive range of motion of the right upper extremity was normative.

An electrodiagnostic examination was performed during a second visit, 4 weeks after the onset of symptoms. By this time her strength had improved slightly, but her sensation disturbance had not changed. Sensory nerve conduction studies showed increased distal latency and 42% reduced amplitude of the right lateral anterbrachial cutaneous nerve when compared to the left. Electromyography of the right biceps and brachialis muscles revealed a pattern of active, ongoing denervation, consisting of fibrillation potentials and positive sharp waves. Decreased recruitment and decreased activation were seen in the right brachialis muscle. The coracobrachialis muscle was normative with needle exam. The remainder of electrodiagnostic testing was also normative (tables 1, 2).

Treatment consisted of resting the affected arm. Follow-up interview by telephone at 5 months revealed complete motor and sensory recovery. The patient returned to skydiving.

**DISCUSSION**

Skydiving and parachuting are inherently dangerous activities, and practitioners of sports medicine should be aware of
injuries that may result. In a study of 110,000 recreational jumps, the rate of injuries requiring medical attention was .14%; rate of fatality was .005%.3 Simulated skydiving activities may decrease danger, and they are less expensive, are easier to repeat, and are therefore useful for training. These include the suspended parachute harness, parachute landing fall practice, and the vertical wind-tunnel, to name a few. For example, the vertically oriented wind-tunnel simulates freefall by producing upward windspeed of nearly 160kph (100mph). In fact, this very realistic freefall simulation is so thrilling that the vertical wind-tunnel is now available as a recreational activity in itself (fig 1). Although simulations curtail some dangers of actual skydiving, however, activities like the vertical wind-tunnel still carry risk of injury. There is only one previous report of peripheral nerve injury related to skydiving—a case of bilateral arm weakness in a recreational parachutist, later discovered to have hereditary neuropathy with predisposition to pressure palsies.3

Although isolated nerve injuries are uncommon, the activities and physiologic demands of athletes can result in these unusual injuries. It is important to be aware of such peripheral nerve injuries to facilitate proper diagnosis and management. Isolated musculocutaneous nerve injury is particularly rare, but many of the reported cases have been sports-related: weightlifting,4,5 resistive exercise,6 rowing,7 football throwing,8 sports-related trauma,9 swimming, tennis,10 racquetball,9 windsurfing,11 and other strenuous sports7 have been implicated. The fibers of the musculocutaneous nerve arise from C5, C6, and sometimes C7, and go through the upper trunk, anterior division, and lateral cord of the brachial plexus.12,13 The musculocutaneous nerve supplies the coracobrachialis, biceps, and brachialis muscles, and terminates as the lateral antebrachial cutaneous nerve, which provides cutaneous sensation to the lateral aspect of the forearm. Musculocutaneous nerve injury in athletes can be mistaken for cervical radiculopathy, brachial plexus injury, median or radial nerve injury, lateral epicondylitis or biceps tendon rupture. For example, Braddock and Wolf2 reported a series including 2 weightlifters who presented with painless weakness in the dominant arm. Both were ultimately diagnosed with musculocutaneous nerve injury. One had been initially misdiagnosed with cervical radiculopathy, however. The other patient had been misdiagnosed with a biceps tendon rupture, and this error was not discovered until unnecessary surgery revealed the intact tendon.2

Proposed mechanisms of musculocutaneous nerve palsy correspond with 2 patterns: proximal injury, resulting in motor and sensory deficits; or distal injury, with primarily sensory deficits. Proximal injury to the musculocutaneous nerve causes a painless syndrome of weakness in the biceps and numbness in the lateral forearm. This may be caused by compression within the coracobrachialis muscle or compression of the muscle or nerve by the humeral head. Distal injury to the musculocutaneous nerve can occur near the bicipital aponeurosis, affecting only the sensory branch, resulting in a painful, pure sensory syndrome.5,13 Reported treatments for musculocutaneous nerve entrapment include rest, ice, nonsteroidal anti-inflammatory drugs, splinting, corticosteroid injection, and surgery.2,13

CONCLUSIONS

In the recreational skydiver described above, symptoms and physical and electrodiagnostic findings suggest an isolated, proximal lesion of the right musculocutaneous nerve, with motor and sensory involvement. Weakness with elbow flexion and forearm supination, and a distinct patch of numbness on the lateral forearm, correspond to the distribution of the musculocutaneous nerve. Even with loss of the biceps and brachialis, the patient was still able to flex her right elbow because the brachioradialis and pronator teres, innervated by intact radial and median nerves, also provide elbow flexion. She was still able to supinate her right forearm with loss of biceps because the brachioradialis and supinator muscles, innervated by the intact radial nerve, also provide supination. Electrodiagnostic studies (see tables 1, 2) further confirmed that the injury was isolated to a single peripheral nerve, because no abnormalities were evident in median, ulnar, radial, or axillary nerves, thus ruling out brachial plexus lesion. A mononeuropathy pattern can be seen in neuralgic amyotrophy, but it does not typically involve the musculocutaneous nerve. Lateral antebrachial cutaneous nerve conduction abnormalities are common in neuralgic amyotrophy due to involvement of the brachial plexus. Specifically, in spite of lateral antebrachial cutaneous nerve conduction abnormalities, neuralgic amyotrophy was not likely

<table>
<thead>
<tr>
<th>Nerve Conduction</th>
<th>Recording Site</th>
<th>Stimulation Site</th>
<th>Latency [ms]</th>
<th>Amplitude</th>
<th>Distance [cm]</th>
<th>Conduction Velocity [m/s]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Right median motor</td>
<td>APB</td>
<td>Wrist</td>
<td>3.40</td>
<td>4.3mV</td>
<td>7</td>
<td>NA</td>
</tr>
<tr>
<td>Right median motor</td>
<td>Antecubital</td>
<td>Wrist</td>
<td>6.90</td>
<td>4.8mV</td>
<td>22</td>
<td>62.9</td>
</tr>
<tr>
<td>Right ulnar motor</td>
<td>ADM</td>
<td>Wrist</td>
<td>2.65</td>
<td>7.8mV</td>
<td>7</td>
<td>NA</td>
</tr>
<tr>
<td>Right ulnar motor</td>
<td>Above elbow</td>
<td>Wrist</td>
<td>5.95</td>
<td>7.5mV</td>
<td>21.5</td>
<td>65.2</td>
</tr>
<tr>
<td>Right ulnar motor</td>
<td>Below elbow</td>
<td>Wrist</td>
<td>7.60</td>
<td>7.3mV</td>
<td>11</td>
<td>66.7</td>
</tr>
<tr>
<td>Right median sensory</td>
<td>Index finger</td>
<td>Forearm</td>
<td>3.10</td>
<td>40.4µV</td>
<td>13</td>
<td>NA</td>
</tr>
<tr>
<td>Right ulnar sensory</td>
<td>Small finger</td>
<td>Forearm</td>
<td>2.90</td>
<td>33.0µV</td>
<td>11</td>
<td>NA</td>
</tr>
<tr>
<td>Right radial sensory</td>
<td>Thumb</td>
<td>Forearm</td>
<td>2.40</td>
<td>12.8µV</td>
<td>10</td>
<td>NA</td>
</tr>
<tr>
<td>Right lateral antebrachial cutaneous</td>
<td>Forearm</td>
<td>Forearm</td>
<td>2.60</td>
<td>11.8µV</td>
<td>12</td>
<td>NA</td>
</tr>
<tr>
<td>Left lateral antebrachial cutaneous</td>
<td>Forearm</td>
<td>Forearm</td>
<td>2.35</td>
<td>20.2µV</td>
<td>12</td>
<td>NA</td>
</tr>
</tbody>
</table>

Abbreviations: ADM, abductor digiti minimi; APB, abductor pollicis brevis; NA, not available.
in our patient because of the absence of the pathognomonic history of pain followed by weakness. Normal electromyography of some C5 and C6 muscles (see table 1) revealed that the injury pattern was not isolated to a spinal segment, as could occur in radiculopathy or plexopathy. These studies were performed about 4 weeks after the injury, which was within the appropriate time window to see evidence of denervation after Wallerian degeneration resulting from axonal disruption (axonotmesis).

Injury was distal to the branches supplying the coracobrachialis muscle, because there was no sign of weakness or denervation in this muscle. Because the musculocutaneous nerve branches and supplies the coracobrachialis before it pierces the muscle, it is possible to injure the nerve within this muscle and spare its innervation. Extension of the shoulder places stretch on the coracobrachialis muscle, which may cause injury to the musculocutaneous nerve. We propose that the mechanism of injury in our patient must have been the abducted, extended, and externally rotated position of the arm during simulated freefall, which caused repeated, prolonged compression of the nerve.

References
**BRIEF REPORT**

**Ultrasonographic Findings of the Normal Ulnar Nerve in Adults**

Michael S. Cartwright, MD, Hae W. Shin, MD, Leah V. Passmore, MS, Francis O. Walker, MD


**Objective:** To provide a detailed description of the ultrasonographic findings along the entire length of the normative ulnar nerve.

**Design:** Volunteers were recruited to undergo ultrasonography of both upper extremities. Age, sex, height, weight, body mass index, arm length, and hand length were recorded, and cross-sectional measurements of the ulnar nerve were obtained at 7 predetermined sites.

**Setting:** The diagnostic neurology laboratory of a referral medical center.

**Participants:** Thirty volunteers (60 arms) were recruited. Volunteers were screened by history and physical examination, and those with evidence of peripheral nervous system disease were excluded.

**Interventions:** Not applicable.

**Main Outcome Measure:** The average cross-sectional area (CSA) of the ulnar nerve at 7 predetermined sites along the entire course of the nerve.

**Results:** The following average ulnar nerve CSAs were obtained: distal wrist crease, 5.9mm²; arterial split, 6.3mm²; 2cm distal to tip of the medial epicondyle, 6.4mm²; tip of the median epicondyle, 6.5mm²; 2cm proximal to tip of the median epicondyle, 6.7mm²; mid-humerus, 6.1mm²; and axilla, 6.2mm². There was no statistical difference in nerve size when dominant and nondominant arms were compared, but women did have smaller nerves than men. Of all the variables measured, nerve size correlated most closely with weight, with a correlation coefficient of .59.

**Conclusions:** The ulnar nerve was easily visualized and measured along its entire course, and the CSA of the nerve was consistent at multiple sites. The reference values obtained in this study will facilitate the analysis of abnormal conditions.

**Key Words:** Body mass index; Nerve entrapments; Rehabilitation; Ulnar nerve; Ultrasonography.

© 2007 by the American Congress of Rehabilitation Medicine and the American Academy of Physical Medicine and Rehabilitation

**HIGH-RESOLUTION ULTRASONOGRAPHY** is an emerging technology for the evaluation of the peripheral nervous system. Over the past 15 years many studies have examined the utility of this technology, but most reports have focused on imaging the median nerve. Some studies have explored imaging of the ulnar nerve, but a detailed evaluation along the entire course of the ulnar nerve has not been performed. The goal of this study was to characterize the ultrasonographic features of the normative ulnar nerve in adults.

An enlarged cross-sectional area (CSA) of the median nerve is a key finding in entrapment neuropathy. Therefore, it was felt that measuring the CSA at several sites along the entire length of the normative ulnar nerve would be of interest and helpful in the evaluation of abnormal conditions.

**METHODS**

After obtaining institutional review board approval, we recruited 30 volunteers (60 arms) from patients and employees in our electrophysiology laboratory over several months in 2005, and we made an effort to select a diverse group of subjects in regard to sex, age, and race. A brief history was obtained from all volunteers, and those with symptoms referable to the peripheral nervous system were excluded. Age, sex, handedness, height, weight, arm length (shoulder to tip of third digit), and hand length (distal wrist crease to tip of third digit) were obtained, and body mass index (BMI) was calculated for each subject.

We performed ultrasonography with a Philips HDI 5000 scanner and 15MHz transducer, and measurements were collected by several operators with training in neuromuscular ultrasound. Reliability testing was not performed, but good interrater reliability for ulnar nerve ultrasound has previously been shown. Seven predetermined sites along the length of the ulnar nerve were imaged. These sites were chosen based on several factors, including common areas of nerve entrapment, sites amenable to reproducible imaging, and sites evenly distributed along the entire course of the nerve. The 7 sites were: (1) distal wrist crease, (2) arterial split (where the nerve separates from the ulnar artery, typically in the mid-forearm), (3) 2cm distal to tip of the medial epicondyle, (4) tip of the medial epicondyle, (5) 2cm proximal to tip of the medial epicondyle, (6) mid-humerus (mid-point between elbow crease and axilla), and (7) axilla. Ultrasonography was performed with the patient supine and the arm abducted 90° from the body. The arm was slightly bent (15°) to facilitate imaging around the elbow. At each site the CSA of the ulnar nerve was obtained by circumferentially tracing just inside the hyperechoic rim of the nerve, and care was made to ensure the transducer was perpendicular to the nerve so the smallest and most accurate CSA was obtained.

We compared nerve size in dominant and nondominant arms, as well as between men and women. Correlation coefficients were calculated by comparing nerve size with height, weight, BMI, arm length, and hand length. Observations were made of nerve shape, echogenicity, location, and relationship.
Fig 1. The CSA of the ulnar nerve at 7 sites is depicted. NOTE. Values are mean mm² ± SD. Next to the measurements are the corresponding ultrasonographic pictures. In these sites the top corresponds to the skin surface and the left represents the right side of the body, similar to the standard viewing of an axial computed tomography. The arrow in each picture indicates the ulnar nerve. At the 2 most distal sites the ulnar artery (UA) can be seen, and at the most proximal site the brachial artery (BA) is adjacent to the ulnar and median (MN) nerves. The ulna (U), humerus (H), medial epicondyle (ME), flexor digitorum profundus (FDP), flexor digitorum superficialis (FDS), flexor carpi ulnaris (FCU), and brachialis (BR) are also depicted. These images were obtained with a Philips iU22 scanner.

### RESULTS

The mean age of the 30 volunteers was 31 years old (range, 24–50y). Nineteen were women, 29 were right handed, mean weight was 71kg, mean height was 166cm, mean arm length was 73cm, and mean hand length was 18cm. Post hoc analysis was performed to determine adequacy of the sample size. We used a computer-intensive bootstrapping technique to estimate the sampling distribution of the variance at each site, and a sample size of 30 adequately estimated the variance. An increase in sample size did not alter the variance estimates, although it would decrease the range of the confidence interval (CI) for the coefficient of variation.

The following mean ulnar nerve CSAs ± standard deviation (SD) were obtained: distal wrist crease, 5.9 ± 1.1mm²; arterial split, 6.3 ± 1.0mm²; 2cm distal to tip of the medial epicondyle, 6.4 ± 1.1mm²; tip of the medial epicondyle, 6.5 ± 0.9mm²; 2cm proximal to tip of the medial epicondyle, 6.7 ± 1.1mm²; mid-humerus, 6.1 ± 0.9mm²; and axilla, 6.2 ± 1.1mm². This information is depicted in figure 1. There was no statistical difference in nerve size when comparing dominant and nondominant arms (P range, .08 –.94, depending on site of comparison), and further side-to-side comparisons with 99% CIs are provided in table 1. Site-to-site variance was measured in the dominant arm only, and no significant differences were noted (P=.14). There were statistically significant differences based on sex at some sites (table 2). Of all variables measured (height, weight, BMI, arm length, hand length), nerve size correlated most closely with weight, with a correlation coefficient of .59.

### DISCUSSION

The ulnar nerve was visible throughout the entire upper extremity and easily identified and measured at all sites. The area of the nerve was consistent throughout its entire length (5.9 to 6.7mm²), and the SD at each site ranged from 0.9 to 1.1mm². These findings are similar to previously published values. Other researchers have used diameter, rather than CSA, when evaluating nerve size, and some have used both measurements. Given the observed variation in the shape of the nerve, from a flat oval at the wrist to a circle at the mid-humerus level (see fig 1), we feel the CSA is a more consistent measurement. CSA is also the measurement of choice for evaluating the median nerve. Standardizing neuromuscular ultrasound measurement techniques will allow comparison between different studies.

### Table 1: Side-to-Side Variation in Ulnar Nerve Size

<table>
<thead>
<tr>
<th>Site</th>
<th>Dominant Arm</th>
<th>Nondominant Arm</th>
<th>Paired Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean ± SD</td>
<td>99% CI</td>
<td>Mean ± SD</td>
</tr>
<tr>
<td>Wrist crease</td>
<td>5.9 ± 1.1</td>
<td>5.4 – 6.4</td>
<td>5.9 ± 1.5</td>
</tr>
<tr>
<td>Arterial split</td>
<td>6.3 ± 1.0</td>
<td>5.8 – 6.7</td>
<td>6.3 ± 1.2</td>
</tr>
<tr>
<td>2cm distal</td>
<td>6.4 ± 1.1</td>
<td>5.9 – 6.9</td>
<td>6.3 ± 1.2</td>
</tr>
<tr>
<td>Epicondyle</td>
<td>6.6 ± 1.1</td>
<td>6.1 – 7.1</td>
<td>6.3 ± 0.9</td>
</tr>
<tr>
<td>2cm proximal</td>
<td>6.7 ± 1.3</td>
<td>6.1 – 7.3</td>
<td>6.7 ± 1.0</td>
</tr>
<tr>
<td>Mid-humerus</td>
<td>6.3 ± 1.0</td>
<td>5.8 – 6.7</td>
<td>6.0 ± 1.0</td>
</tr>
<tr>
<td>Axilla</td>
<td>6.2 ± 1.2</td>
<td>5.7 – 6.8</td>
<td>6.2 ± 1.1</td>
</tr>
</tbody>
</table>

NOTE. Values are in mm². This table presents the 7 ulnar nerve CSA measurements for the dominant and nondominant arms. No statistically significant differences were found with side-to-side comparisons. Normal variation was defined by the 99% CI for paired difference.
Table 2: Sex Differences in Ulnar Nerve Size

<table>
<thead>
<tr>
<th>Site</th>
<th>Men (n=11) Mean ± SD</th>
<th>Women (n=19) Mean ± SD</th>
<th>99% CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wrist crease</td>
<td>6.4±1.2</td>
<td>5.4–7.4</td>
<td>5.6±0.8</td>
<td>.049</td>
</tr>
<tr>
<td>Arterial split</td>
<td>6.5±1.2</td>
<td>5.6–7.5</td>
<td>5.7±0.7</td>
<td>.249</td>
</tr>
<tr>
<td>2cm distal</td>
<td>6.9±1.4</td>
<td>5.8–8.0</td>
<td>5.7±0.8</td>
<td>.065</td>
</tr>
<tr>
<td>Epicondyle</td>
<td>7.2±1.3</td>
<td>6.1–8.2</td>
<td>5.8±0.8</td>
<td>.020</td>
</tr>
<tr>
<td>2cm proximal</td>
<td>7.2±1.1</td>
<td>6.4–8.1</td>
<td>5.6±1.3</td>
<td>.072</td>
</tr>
<tr>
<td>Mid-humerus</td>
<td>6.8±1.3</td>
<td>5.8–7.8</td>
<td>5.9±0.7</td>
<td>.019</td>
</tr>
<tr>
<td>Axilla</td>
<td>6.8±1.7</td>
<td>5.5–8.1</td>
<td>5.9±0.7</td>
<td>.066</td>
</tr>
</tbody>
</table>

NOTE. Values are in mm². This table presents the 7 ulnar nerve CSA measurements for men and women. Three sites (wrist crease, epicondyle, mid-humerus) show statistically significant differences (P<.05).

In addition to noting variation in nerve shape, we made several other observations during this study. First, the nerve is superficial throughout the entire arm, which makes it easy to visualize. Second, although no measurements were obtained distal to the wrist crease, the ulnar nerve can be followed several centimeters into the palm without difficulty. Once the nerve begins to divide, however, the distal branches are difficult to track with ultrasound. Third, the entire ulnar nerve is often hypoechoic in the ulnar groove, and discrete fascicles are not easily seen. Knowledge of this finding may help when evaluating ulnar nerve pathology in the cubital tunnel with ultrasound. Fourth, subluxation of the ulnar nerve, a finding previously described with ultrasound, was noted in several subjects. Finally, during several studies pulsatile vessels were seen accompanying the nerve within the ulnar groove. Further study into the significance of this finding is warranted, because these vessels could theoretically predispose to entrapment and also may be of interest when planning surgical intervention for ulnar neuropathy at the elbow.

Side-to-side comparisons are used in the electrodiagnostic laboratory because they may be more sensitive for detecting abnormalities than comparing results to reference values alone. This study showed no statistical difference in nerve CSA between dominant and nondominant arms, but perhaps more important, normal side-to-side variability was defined (see table 1).

At 3 of the 7 sites we saw statistically significant differences between men and women, and the general trend was for women to have smaller ulnar nerve CSAs. Normative measurements by gender are reported in table 2, and this sex-specific information should help further define normal and abnormal states.

The CSA of the ulnar nerve was compared with 5 different body habitus measurements. In a previous study, ulnar nerve size correlated significantly with BMI. We too found a correlation with BMI, but we found that ulnar nerve size correlated most closely with weight. Knowledge of this relationship may be useful for interpreting borderline nerve enlargement in obese people, although further research into this relationship is needed because other studies have failed to show a correlation between the size of other nerves and BMI.

CONCLUSIONS

This is the first study to report reference values for the CSA along the entire course of the ulnar nerve. The nerve was easily visualized and measured along its entire length, and the CSA of the nerve was consistent at multiple sites. Nerve size correlated most closely with weight. We feel the reference values obtained in this study are essential for facilitating the analysis of abnormal nerve conditions, and the information on side-to-side variation and sex-specific differences will be particularly helpful. These values can be used for comparison when examining different disease states, such as entrapments, trauma, hereditary neuropathies, and degenerative nerve diseases.

References

Supplier
a. Philips Medical Systems, 22100 Bothell Everett Hwy, Bothell, WA 98041-3003.
Letter to the Editor

Exercise Interventions for Diabetes Control: Do We Really Know That Strength Training Is Better Than Endurance Training?

Recently Cauza et al.1 compared the influence of strength training to endurance training on glycemic control, atherogenic markers, cardiovascular risk factors, body composition, strength, and cardiorespiratory endurance in people with diabetes. At first glance, results seemed clearly definitive that strength training was superior to endurance training. However, on closer inspection, the analysis strategy proposed in the methods section did not appear to be the analysis on which reported results were based. The risk in this is that conclusions of strength training superiority may be based on analyses that are insufficient to establish superiority.

The methods section indicated that 2-factor analysis of variance (ANOVA) was to be employed. Although not specified, presumably this was to be a mixed-model ANOVA comparing groups (strength vs endurance) by time of measurement (before vs after training). This would have been an appropriate analysis strategy, assuming that groups did not differ on relevant variables before training. Unfortunately, it appears that the analysis strategy employed neither followed this plan nor accounted for differences among groups before training on such important variables as plasma glucose levels and total cholesterol. The strength group had significantly poorer scores on each of these measures before training, allowing for a greater margin for change in post-training scores due to regression to the mean in combination with any effect of strength training. In fact, table 2 reveals strikingly poorer pretraining scores for the strength training group on blood glucose, plasma insulin, glycosylated hemoglobin (Hb A1c), insulin resistance, high-density and low-density lipoprotein cholesterol, and total cholesterol. A more appropriate analysis would have employed “before training” scores as a covariate in an analysis seeking differences among post-training means after adjustment for pretraining group differences. Analysis of covariance has long been recommended as a method to reduce bias when groups differ at baseline.2

Inferential decisions were based on analyses that are insufficient to establish superiority. The significance of treatment effects did not appear to be the analysis on which reported results were based. The method to analyze data is less aggressive than strength training. Endurance training with our intensity and volume in this study was less aggressive than strength training. Because exercise is known to improve health in these patients, they felt that it would be unethical not to recommend exercise. Perhaps endurance training with our volume and intensity in this study was less aggressive than strength training. Endurance training with an intensity of 60% of maximal oxygen consumption and a low (frequency) volume (starting at 15 min and advancing to a maximum of 30 min 3 times a week) was at the time.

What is the risk in basing conclusions on analyses that, at best, are not adequately described and, at worst, result in analyses that favor one group over the other? Clinicians depend on the results of well-analyzed trials as evidence for patient care. I commend Cauza for using randomized methodology to study 2 clinically important interventions head-to-head. However, it remains unclear whether the differences in favor of strength training would exist when an analysis comparing groups simultaneously over time was employed. Without a more precise analysis, it would be premature for clinicians to recommend strength training over endurance training for patients with type 2 diabetes.

Douglas L. Weeks, PhD
St Luke’s Rehabilitation Institute
Spokane WA
weeksl@inhs.org

References


The author responds

I would like to address the significant concerns raised by Douglas Weeks in relation to our study.1

Because endurance training has been advocated as the most suitable exercise mode in type 2 diabetes (T2D), my colleagues and I wanted to compare the effects of strength training with the most effective exercise mode. Therefore, we used endurance training for the control group. We also intended to compare the effects of strength training with a second control group, a group without training, but our ethics committee had reservations about this approach. Because exercise is known to improve health in these patients, they felt that it would be unethical not to recommend exercise.

Perhaps endurance training with our intensity and volume in this study was less aggressive than strength training. Endurance training with an intensity of 60% of maximal oxygen consumption and a low (frequency) volume (starting at 15 min and advancing to a maximum of 30 min 3 times a week) was at the time.

doi:10.1016/j.apmr.2007.01.008

lower end of an endurance training program. We opted for this approach because of the extremely poor physical condition of our patients at study entry. All patients were untrained, so we had to start with low intensity and low volume.

The main aim of our study was to demonstrate that strength training is effective in T2D patients. We were able to show that strength training leads to benefits in glycemic and metabolic parameters and showed that strength training is neither dangerous nor ineffective. However, we agree with Weeks in that strength training in this study was not superior to endurance training but at least led to equal benefit.

Although all patients were carefully and correctly randomized, fasting blood glucose (FBG) and triglyceride values at study entry were significantly increased and glycosylated hemoglobin (Hb A1c) slightly (but insignificantly) increased in the strength training group.

All other parameters (including plasma insulin, all cholesterol values, blood pressure, and several measurements of body composition) did not differ significantly at study entry. It is known that values that are highly elevated and outside of a normative physiologic range can be reduced more easily than values that are closer to the normative range. Although the Hb A1c and FBG levels in the strength training group were higher and outside of physiologic range at study entry, the values were closer to normative levels than in the endurance training group after the 4-month training period. We observed similar effects for triglycerides, with nearly equal values in both groups after the training period.

The statistical analysis plan was developed by Heinz Tüchler, a statistician at the Ludwig Boltzmann Institute for Leukaemia Research and Haematology, before the study started. After consulting with him, multiple t tests were considered appropriate for this study design. In addition to the selected method, a statistical analysis using analysis of variance (ANOVA) was also identified as applicable to our study. Consequently, we changed the statistical evaluations and used ANOVA to examine differences in all parameters. The results of the ANOVA analysis were identical to the results of the t tests. Because the goal of our study was to show that strength training may be effective in T2D patients, our statistican used ANOVA to analyze the data.

Edmund Cauza, MD, PhD
Department of Diabetes and Metabolism
Wilhelminenspital
Vienna, Austria

e-mail: edmund.cauza@wienkav.at

Reference

doi:10.1016/j.apmr.2007.01.020


DEPARTMENTS

New Books

Brain Injuries

Communication Disorders

General PM&R

Neurologic Disorders

Neuropsychology

Orthotic Devices

Pharmacology

Vascular Disease

Vision Rehabilitation
Corrections

The following 3 abstracts were marked as “Canceled” and not included in the online edition of the October 2006 issue:

Poster 11
The Acute Impact of Stroke on Functioning. Lois E. Finch, MSc (McGill University, Montreal, QC, Canada) J.M. Higgins, N.E. Mayo.

Disclosure: None declared.

Objective: To develop a measure of the acute impact of stroke on functioning by combining items from self-report and observational indices.

Design: Observational. Setting: A large urban acute care hospital stroke unit. Participants: Consecutive subjects (N=235) admitted within 3 days of a stroke, either ischemic or hemorrhagic. Subjects were predominately male (age ± SD, 71.6±12.5y), with ischemic stroke. Interventions: Not applicable. Main Outcome Measures: Functional assessments were made at 3 days poststroke, using 11 indices and tests commonly used to evaluate stroke survivors (total items, 171). Information on potentially important variables was also collected. Rasch analysis assisted in constructing the calibrated linear measure. Results: The 171 items reduced to a 38-item unidimensional measure, the acute impact of stroke on functioning, containing items on upper- and lower-extremity movement, balance, mobility, and activities of daily living. All Rasch model assumptions were met. Both person and item reliability (.97) indicated a stable hierarchy. Precision of the items and persons were .37 and 1.3 logits, respectively. Conclusions: A single index quantifying the acute impact of stroke on functioning was developed. This measure may assist in matching patients to therapeutic options earlier. Key Words: Recovery of function; Rehabilitation.

Poster 15
Safety and Efficacy of Repeated Botulinum Toxin Type A Treatments for Focal Upper-Limb Poststroke Spasticity: Results of a 12-Month Multicenter, Open-Label Trial. E. Elovic (Kessler Medical Rehabilitation Research and Education Corp, West Orange, NJ), A. Brashear, D. Kaelin, R. McIntosh, J. Liu, C. Turkel.

Disclosure: Supported by Allergan. Elovic, Brashear, and Kaelin have received grant support from Allergan and have been or are on Allergan’s speakers bureau; Liu and Turkel are employees of Allergan.

Objective: To evaluate long-term safety and efficacy of repeated botulinum toxin type A (BTX-A) (Botox) injections in the treatment of focal upper-limb poststroke spasticity. Design: Multicenter, open-label study. Setting: 35 North American study centers. Participants: 279 upper-limb poststroke spasticity patients. Interventions: Up to 5 BTX-A treatments (200–400U each) administered to the wrist, fingers, and elbow flexors. Main Outcome Measures: Safety and efficacy were assessed by reported adverse events (AEs), the Ashworth Scale, and the Disability Assessment Scale. One disability domain was selected and measured as the principal therapeutic intervention target. Results: 6.5% of patients reported AEs judged to be treatment-related, none of which was serious. Muscle tone in the wrist, fingers, and elbow improved considerably by week 6 and functional disability also improved with at least 50% of patients achieving a minimum 1-point improvement in principal therapeutic intervention target. Conclusions: Repeated BTX-A treatment safely improves upper-limb spasticity in poststroke patients. The low incidence of AEs represents an advantage over oral antispasticity medications associated with systemic side effects. Key Words: Botulinum toxin type A; Muscle spasticity; Rehabilitation.


ARTICLE OBJECTIVES:

Article One: Effect of Motorized Scooters on Physical Performance and Mobility: A Randomized Clinical Trial

Learning Objective: After completing this article and with appropriate self-study, the participant will be able to:
   a) Discuss the effect of scooter usage on the 6-minute walk test.
   b) List 3 places people use scooters. Discuss their importance.
   c) Recall the accident rate for scooter users and the most common type of accident.

Article Two: Spasticity Experience Domains in Persons With Spinal Cord Injury

Learning Objective: After completing this article and with appropriate self-study, the participant will be able to:
   a) Recall 2 symptoms that patients with spinal cord injury may relate when asked about spasticity.
   b) List 3 subcategories related to the physical nature of spasticity.
   c) List 3 issues related to functional self-management of spasticity.

Article Three: Motor Points for the Neuromuscular Blockade of the Subscapularis Muscle

Learning Objective: After completing this article and with appropriate self-study, the participant will be able to:
   a) Recall the function of the subscapularis muscle in patients with spasticity.
   b) Describe the “best fit” line of motor points of the subscapularis muscle.
   c) Describe the technique for injecting botulinum toxin type A into this area.

INDICATE THE DEGREE TO WHICH YOU AGREE OR DISAGREE WITH EACH STATEMENT

1. Objectives for article one were met.  SD.............D.............NC.............A.............SA
2. Objectives for article two were met.  SD.............D.............NC.............A.............SA
3. Objectives for article three were met.  SD.............D.............NC.............A.............SA
4. I learned a new skill or patient management approach.  SD.............D.............NC.............A.............SA
5. This material will enhance my professional effectiveness.  SD.............D.............NC.............A.............SA
6. I plan to implement a change(s) to my practice as a result of this material.  SD.............D.............NC.............A.............SA

   If you circled “Agree” or “Strongly Agree,” please give one example:

___________________________________________________________________________
___________________________________________________________________________

7. In what ways did/will you utilize the information from these articles in your medical practice? I have used/will use it to:
   (Check all that apply.)
   □ Confirm previous knowledge and reinforce clinical practice
   □ Study for recertification examination
   □ Serve as initial resource for clinical problems
   □ Apply new techniques/procedures to the care of my patients
   □ Use the information to train staff
   □ Share the information with colleagues
   □ Help develop new policy and procedures
   □ Other (please specify):

8. The material was fair, objective, and unbiased toward any product or program.................................Yes .......No

9. Please share any general comments, recommendations, or an elaboration of any item on this form:

___________________________________________________________________________
___________________________________________________________________________

Evaluation data collected from this form will be processed confidentially by a third party and will only be reviewed by Academy staff in the aggregate.
CME PROCESSING FEES
AND
APPLICATION

CME Certificates will not be processed without the completion of the Evaluation

<table>
<thead>
<tr>
<th>Membership/ID Number</th>
<th>Fee</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAPM&amp;R/ACRM member</td>
<td>$20.00</td>
</tr>
<tr>
<td>Non-member</td>
<td>$25.00</td>
</tr>
</tbody>
</table>

TO RECEIVE CME CREDIT: YOU MUST COMPLETE THE FOLLOWING 5 ITEMS.

The American Academy of Physical Medicine and Rehabilitation is accredited by the Accreditation Council for Continuing Medical Education to provide continuing medical education for physicians. The AAPM&R designates this educational activity for a maximum of 3 *AMA PRA Category 1 Credits™*. Physicians should only claim credit commensurate with the extent of their participation in the activity.

1. Membership/ID Number: ____________________________
2. AAPM&R Member ____  ACRM Member ____  Non Member____
3. Enclose relevant Program Evaluation(s).
4. Number of hours:
   Credit may be obtained for up to 3 years from date of publication.

☐ I studied the articles identified as eligible for CME in this issue of the *Archives of Physical Medicine and Rehabilitation*, consulted the reference materials as needed, and completed the corresponding evaluation.

Therefore, I claim ____ (fill in blank) *AMA PRA Category 1 Credits™* (up to 3 hours).

5. To ensure that we accurately process your CME certificate, please print/type your full name, degree(s), and address as they should appear on your certificate.

Name and Degrees: ____________________________________________
Address: ________________________________________________
_________________________ _____________________________
_________________________ _____________________________
_________________________ _____________________________

Signature: _____________________________ Date: ______________

Your CME Certificate will be mailed to you following receipt of your application, evaluation, and check in the Academy Office.

PLEASE COMPLETE AND RETURN THIS CME APPLICATION, EVALUATION AND PAYMENT TO:
AAPM&R
PO Box 95528
Chicago, IL 60694-5528
AAPM&R 2007 Study Guide Included With This Issue of Archives

The AAPM&R 2007 Study Guide/Self-Assessment for Practitioners (SAE-P) (print and online versions) is included as a supplement to the March issue of Archives of Physical Medicine and Rehabilitation. This edition covers 2 key clinical topic areas:

- Industrial Rehabilitation Medicine and Acute Musculoskeletal Rehabilitation
- Spinal Cord Injury Medicine

The AAPM&R Study Guide may be used as a means to fulfill the Lifelong-Learning and Self-Assessment component for the Maintenance of Certification program offered by the American Board of Physical Medicine and Rehabilitation.

The Study Guide/SAE-P is a continuing medical education (CME) tool designed to guide both physiatrists and residents in physical medicine and rehabilitation in their self-directed study.

AAPM&R is accredited by the Accreditation Council for Continuing Medical Education to provide CME for physicians. AAPM&R designates this educational activity for a maximum of 30.0 AMA PRA Category 1 Credit(s)™. Physicians should only claim credit commensurate with the extent of their participation in the activity.

AMA PRA Category 1 Credit(s)™ may be obtained for up to 3 years from the date of publication.

For questions regarding the Study Guide, contact Duane Kinoshita at the Academy national office at 312-464-9700 or dkinoshita@aapmr.org. For questions about the online access of the Study Guide, contact Elsevier at 800-654-2452 (USA, Canada) or elsols@elsevier.com or eurorsupport@elsevier.com.

2007 AAPM&R Officers Elected

Joel M. Press, MD, assumed the presidency of the AAPM&R during the 2006 Annual Assembly in Honolulu, HI, November 9–11, 2006. Dr. Press focuses on the care of persons with musculoskeletal, spine, and sports-related disorders. An assistant professor of physical medicine and rehabilitation, Northwestern University Medical School, Dr. Press is board certified in both PM&R and electrodiagnostic medicine. He is an attending physician at the Rehabilitation Institute of Chicago, where he founded the sports rehabilitation program in 1989 and has served as medical director for the Spine & Sports Rehabilitation Center since 1994.

Dr. Press has served on the AAPM&R Board of Governors since 1999. A founding member and past president of PASSOR, he is also a past president of the North American Spine Society. Dr. Press is a widely published writer and well-known lecturer on PM&R topics.

David S. Cifu, MD, was named president-elect of the Academy. An academic physiatrist, Dr. Cifu is professor and chair, Department of PM&R, Virginia Commonwealth University/Medical College of Virginia. He is also chief of PM&R services for the Virginia Commonwealth University Health System. Dr Cifu will succeed Dr. Press at the completion of the 2007 Annual Assembly in Boston.

William F. Micheo, MD, professor and chair, Department of PM&R, University of Puerto Rico School of Medicine was elected vice president. Kurt Hoppe, MD, a consultant at the Mayo Clinic in Rochester, MN, and assistant medical director for quality and utilization management at Mayo Health Company, was named secretary.

The new AAPM&R Board of Governors will include a new members-at-large, Michael W. O’Dell, MD, associate chief of rehabilitation medicine at New York Presbyterian: The University Hospital of Columbia and Cornell.

Prather Elected New PASSOR President

Heidi Prather, DO, an assistant professor and chief of the section of PM&R in the Department of Orthopaedic Surgery at Washington University School of Medicine in St. Louis, assumed the presidency of the Physiatric Association of Spine, Sports, and Occupational Rehabilitation physicians (PASSOR) during the 2006 PASSOR Business Meeting on November 10, 2006, in Honolulu, HI. Also elected at that meeting were:

Vice President: Sheila Dugan, MD
Secretary/Treasurer: Anne Z. Hoch, DO
Member-at-Large: Venu Akuthota, MD
Member-at-Large: Brian A. Casazza, MD
Member-at-Large: Joseph P. Zuhosky, MD

MUSCULOSKELETAL EDUCATION AND RESEARCH GRANT AVAILABLE FROM FPMR

The Foundation for Physical Medicine and Rehabilitation (FPMR) is now accepting applications for its Musculoskeletal (MSK) Education and Research Grant. One grant (up to $40,000) will be made; the application deadline is March 15, 2007. The MSK Award is a grant for research or education on a topic related to the field of PM&R. Applicants must be members of the AAPM&R and/or the Association of Academic Physiatrists. The project should focus on medical rehabilitation for musculoskeletal conditions. Special consideration will be given to applications focused on outcomes research.

Applications must be submitted on the MSK Education and Research Grant Application Form (which can be found at http://www.foundationforpmr.org/programs/grants.html) and must include a concise statement of the purpose of the project; anticipated outcomes; proposed budget (showing other sources of cash and in-kind support); and a letter of support from the sponsoring institution. Applicants should also include a curriculum vitae. Indirect costs should comprise no more than 15% to 25% of the budget. Applications and all required attachments must be postmarked no later than March 15, 2007.

The MSK Education and Research Grant is funded by interest on an endowment provided by PASSOR along with new donor gifts. Award will be presented at the 2007 AAPM&R Annual Assembly (September 27–30, 2007, in Boston). If you have questions about the award or the application process, please e-mail info@foundationforpmr.org or call 312-464-6646.

AMA and AAPM&R to Conduct Physician Practice Information Survey

The American Medical Association (AMA) with the support of the AAPM&R and more than 60 other medical specialty societies will begin conducting a multispecialty survey of America’s physician practices beginning in 2007. The purpose of the survey is to collect up-to-date information on physician practice characteristics in order to develop and redefine AMA and AAPM&R policy. Data related to professional practice expenses will also be collected. The AMA and AAPM&R will survey thousands of physicians over the year from virtually all physician specialties to ensure accurate and fair representation for all physicians and their patients.

During the year 2007, Academy members may be contacted by the Gallup Organization to participate in this study. The Academy encourages participation in this
survey, as the data obtained will be a critical source of information for the AMA and AAPM&R. Should you be called on to contribute, your participation ensures that the information collected will represent you and your patients’ concerns to national policy-makers. Please watch for this survey in 2007 and do your part in completing it in a thorough and accurate manner.

**Twenty-Five Year Members Honored by Academy**

The Academy recently honored those PM&R physicians who have dedicated themselves to the specialty for a quarter of a century as members of AAPM&R. The growth of PM&R is owed in no small part to these physicians. Please join the Academy in recognizing the following members:

Chang Won Ahn, MD  
Ramana V. Amaravadi, MD  
Carmen A. Angles, MD  
Dennis J. Bonner, MD  
Lila Cherry, MD  
Chin Ting Chiu, MD  
Edward A. Collacott, MD  
Ignatius Cyriac, MD  
Donna D. Davidoff, MD  
Robert W. DePompolo, MD  
Frances G. Deppe, MD  
Armen J. Dumas, MD  
John B. Edwards, MD  
Carl F. Eiben, MD  
Elie P. Elovic, MD  
Richard S. Goka, MD  
Susan K. Halter, MD  
James B. Hoover, MD  
Tibor Ketzan, MD  
C. George Kevorkian, MD  
Choong Kwon Kim, MD  
John J. Kraus, MD, MMM  
Aaron M. Levine, MD  
Ernie Er-Yuh Lin, MD  
Betty Joan Maly, MD  
Philip Messah, MD  
Barry C. Norris, MD  
Lester A. Owens, DO  
Jai Sook Park, MD  
Bernard M. Portner, MD  
Alicia E. Quiroga, MD  
Asit K. Ray, MD  
B.M. Reddy, MD  
Moola Prabhakar Reddy, MD  
Jay D. Roberts, MD  
Mark F. Rottenberg, MD  
Tabassum Saeed, MD  
Han Ki Sim, MD  
J. Craig Stevens, MD  
Benjamin M. Sucher, DO  
Omprakash N. Sureka, MD  
Robert L. Swartzzel Jr, MD  
Gregory M. Thomas, MD  
Sirisha R. Tikko, MD  
Edward A. Trudeau, MD, SC  
Raman L. Vij, MD  
Gary Allen Ward, MD  
Julia D. Weinerman, MD  
Michael S. Weiss, MD  
Gary M. Yarkony, MD  
Chandra Yoganathan, MD  
Kyung Ook Yoon, MD