### Table of Contents

**ORIGINAL ARTICLES**

955  **Circuit Class Therapy Versus Individual Physiotherapy Sessions During Inpatient Stroke Rehabilitation: A Controlled Trial**  
Coralie K. English, PhD, Susan L. Hillier, PhD, Kathy R. Stiller, PhD, Andrea Warden-Flood, PhD

964  **Kinematic and Clinical Analyses of Upper-Extremity Movements After Constraint-Induced Movement Therapy in Patients With Stroke: A Randomized Controlled Trial**  
Ching-yi Wu, ScD, OTR, Chia-ling Chen, MD, PhD, Simon F. Tang, MD, Keh-chung Lin, ScD, OTR, Ya-ying Huang, MS

971  **A Randomized Controlled Trial of an Implantable 2-Channel Peroneal Nerve Stimulator on Walking Speed and Activity in Poststroke Hemiplegia**  
Anke I. Kottink, MSc, Hermie J. Hermens, PhD, Anand V. Nene, MD, PhD, Martin J. Tenniglo, PT, Hans E. van der Au, MD, PhD, Hendrik P. Buschman, PhD, Maarten J. IJzerman, PT, PhD

979  **Staphylococcus aureus Colonization in Community-Dwelling People With Spinal Cord Dysfunction**  
Mary-Claire Roghmann, MD, MS, Peter H. Gorman, MD, Mitchell T. Wallin, MD, Kristen Kreisel, MS, Simone Shurland, MS, Judith A. Johnson, PhD

984  **Role of Sonographic Examination in Traumatic Knee Internal Derangement**  
Chung-Yuan Wang, MD, Hsing-Kuo Wang, PhD, Chao-Yu Hsu, MD, Jeng-Yi Shieh, MD, Tyng-Guey Wang, MD, Ching-Chuan Jiang, PhD

988  **Greater Trochanteric Pain Syndrome: Epidemiology and Associated Factors**  
Neil A. Segal, MD, David T. Felson, MD, James C. Torner, PhD, Yanyan Zhu, MSc, Jeffrey R. Curtis, MD, MPH, Jingbo Niu, DSc, Michael C. Nevitt, PhD, for the Multicenter Osteoarthritis (MOST) Study Group

993  **The Lower-Limb Tasks Questionnaire: An Assessment of Validity, Reliability, Responsiveness, and Minimal Important Differences**  
Peter J. McNair, PhD, Harry Prapavessis, PhD, Jill Collier, MHSc, Sandra Bassett, PhD, Adam Bryant, PhD, Peter Larmer, MHSc

1002  **Quantitative and Clinical Description of Postural Instability in Women With Breast Cancer Treated With Taxane Chemotherapy**  
Meredith A. Wampler, PT, DPTSc, Kimberly S. Topp, PT, PhD, Christine Miaskowski, RN, PhD, Nancy N. Byl, PT, PhD, Hope S. Rugo, MD, Kate Hanel, PhD

1009  **Changes in Postural Control in Hemiplegic Patients After Stroke Performing a Dual Task**  
Laurent Bensoussan, MD, Jean-Michel Viton, MD, PhD, Marco Schieppati, MD, Herve Collado, MD, Virginie Milhe de Bovis, MD, Serge Mesure, PhD, PT, Alain Delarque, MD

1016  **Predictability of Simple Clinical Tests to Identify Shoulder Pain After Stroke**  
Bala S. Rajaratnam, MAppSc (PT), N. Venketasubramanian, FRCP, Prem V. Kumar, FRCS, James C. Goh, PhD, Y.-H. Chan, PhD

1022  **Variation of Muscle Coactivation Patterns in Chronic Stroke During Robot-Assisted Elbow Training**  
Xiaoling Hu, PhD, Kai Y. Tong, PhD, Rong Song, PhD, Vincent S. Tsang, MSc, Penny O. Leung, BEng, Le Li, MSc

1030  **Osteoporosis in a Postpolio Clinic Population**  
Mariel Haziza, MD, Richard Kremer, MD, Andrea Benedetti, PhD, Daria A. Trojan, MD

1036  **Lowering of Sensory, Motor, and Pain-Tolerance Thresholds With Burst Duration Using Kilohertz-Frequency Alternating Current Electric Stimulation**  
Alex R. Ward, PhD, Stacey Lucas-Toumbourou, MPH
Table of Contents (continued)

1042 Pain Perception After Running a 100-Mile Ultramarathon
Martin D. Hoffman, MD, Jean Lee, MS, MD, Holly Zhao, MD, PhD, Alex Tsodikov, PhD

1049 Learning Effects of Repetitive Administrations of the Sensory Organization Test in Healthy Young Adults
Diane M. Wrisley, PhD, PT, NCS, Marilee J. Stephens, PhD, Shaun Mosley, PT, DPT, Adam Wojnowski, AuD, Jordan Duffy, PT, DPT, Robert Burkard, PhD

1055 Strength Training Improves the Tri-Digit Finger-Pinch Force Control of Older Adults
Justin W. Keogh, PhD, Steve Morrison, PhD, Rod Barrett, PhD

REVIEW ARTICLE
1064 What Do We Really Know About the Transition to Adult-Centered Health Care? A Focus on Cerebral Palsy and Spina Bifida
Jessie A. Binks, MSc, Wendy S. Barden, MSc, Tricia A. Burke, BA, Nancy L. Young, PhD

BRIEF REPORT
1074 Energy Cost of the Case Western Reserve Standing Neuroprosthesis
George P. Forrest, MD, Thomas C. Smith, MD, Ronald J. Triolo, PhD, Jason P. Gagnon, PT, Darryl DiRisio, MD, Michael E. Miller, MS, Lori Murray, PT, John A. Davis, MD, Atif Iqbal, MD

COMMENTARIES
1077 Expert Opinion and Controversies in Sports and Musculoskeletal Medicine: Concussion in the Young Athlete
Christopher J. Standaert, MD, Stanley A. Herring, MD, Robert C. Cantu, MD, MA

1080 Perspectives on Transitions: Rethinking Services for Children and Youth With Developmental Disabilities
Peter Rosenbaum, MD, FRCP(C), Debra Stewart, MSc, OT Reg (Ont)

DEPARTMENTS
1083 Letters to the Editor
Consortium for Spinal Cord Medicine

1083 Corrections

ORGANIZATION NEWS—ACRM
1086 Journal-Based CME Evaluation and Application
1088 In Memoriam: Mitchell Rosenthal, PhD (1949-2007)
Official Journal of the American Congress of Rehabilitation Medicine and the American Academy of Physical Medicine and Rehabilitation

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.

EDITOR-IN-CHIEF
Jeffrey R. Basford, MD, PhD (Academy)
Rochester, MN

ACADEMY EDITOR
Leighton Chan, MD, MPH
Bethesda, MD
Managing Editor
Michael A. Vasko, MA
Editorial Assistant
Kristen R. Golem

CONGRESS EDITOR
Allen W. Heinemann, PhD, ABPP
Chicago, IL
Editorial Coordinators
Carol J. Manow
Lupe Soto
Editorial Administrator
Karen K. Parks

Editorial Board
Patrick K. Murray, MD (Academy)
Cleveland, OH
Mary M. Rodgers, PhD, PT (Congress)
Baltimore, MD
Elliot J. Roth, MD (Congress)
Chicago, IL
Dale C. Strasser, MD (Academy)
Atlanta, GA
Denise G. Tate, PhD, ABPP (Congress)
Ann Arbor, MI
Robert A. Werner, MD (Academy)
Ann Arbor, MI

Associate Board
Mark S. Kaplan, MD (Academy)
Boston, MA
Christina M. Marciniak, MD (Academy)
Chicago, IL
Julie D. Moreland, MSc (Congress)
Hamilton, ON
Michael W. O'Dell, MD (Congress)
New York, NY
Jeffrey B. Palmer, MD (Congress)
Baltimore, MD
Thomas D. Rizzo Jr, MD (Academy CME Editor)
Jacksonville, FL
Joan C. Rogers, PhD, OTR/L (Congress)
Pittsburgh, PA
Bruce Shapiro, MD (Academy)
Baltimore, MD
Robert W. Teasell, MD, FRCPC (Academy)
London, ON
Richard D. Zorowitz, MD (Academy)
Baltimore, MD

Editorial Office: Suite 2510, 330 North Wabash Avenue, Chicago, Illinois 60611-7617. Phone: (312) 464-9550. Fax: (312) 464-9554. E-mail: ArchivesMail@aapmr.org Internet Home Page: www.archives-pmr.org
Circuit Class Therapy Versus Individual Physiotherapy Sessions During Inpatient Stroke Rehabilitation: A Controlled Trial

Coralie K. English, PhD, Susan L. Hillier, PhD, Kathy R. Stiller, PhD, Andrea Warden-Flood, PhD†


Objective: To compare the effectiveness of circuit class therapy and individual physiotherapy (PT) sessions in improving walking ability and functional balance for people recovering from stroke.

Design: Nonrandomized, single-blind controlled trial.

Setting: Medical rehabilitation ward of a rehabilitation hospital.

Participants: Sixty-eight persons receiving inpatient rehabilitation after a stroke.

Interventions: Subjects received group circuit class therapy or individual treatment sessions as the sole method of PT service delivery for the duration of their inpatient stay.

Main Outcome Measures: Five-meter walk test (5MWT), two-minute walk test (2MWT), and the Berg Balance Scale (BBS) measured 4 weeks after admission. Secondary outcome measures included the Iowa Level of Assistance Scale, Motor Assessment Scale upper-limb items, and patient satisfaction. Measures were taken on admission and 4 weeks later.

Results: Subjects in both groups showed significant improvements between admission and week 4 in all primary outcome measures. There were no significant between group differences in the primary outcome measures at week 4 (5MWT mean difference, .07m/s; 2MWT mean difference, 1.8m; BBS mean difference, 3.9 points). A significantly higher proportion of subjects in the circuit class therapy group were able to walk independently at discharge (P = .01) and were satisfied with the amount of therapy received (P = .007).

Conclusions: Circuit class therapy appeared as effective as individual PT sessions for this sample of subjects receiving inpatient rehabilitation poststroke. Favorable results for circuit classes in terms of increased walking independence and patient satisfaction suggest this model of service delivery warrants further investigation.

Key Words: Cerebrovascular accident; Physical therapy modalities; Rehabilitation.
balance than individual PT sessions. Our secondary objective was to investigate between-group differences in level of walking independence achieved, upper-limb function, length of stay (LOS) in rehabilitation, and patient satisfaction with therapy.

METHODS

Study Design

We conducted a single-blind, nonrandomized controlled trial. We obtained ethics approval from the University of South Australia and Royal Adelaide Hospital Research Ethics Committees.

Participants

Persons received inpatient rehabilitation poststroke. Specific inclusion and exclusion criteria were as follows.

Inclusion criteria. Subjects who were diagnosed with a cerebrovascular accident resulting in unilateral motor deficits; had sufficient ability to participate in circuit class therapy (ie, ability to follow 3-part commands, sit unsupported and stand with 1 person assisting); and were able to give informed consent.

Exclusion criteria. Persons who had suffered a cerebellar lesion; had a history of any neurologic disorder (excluding previous stroke); or regularly used a walking aid (excluding single-point cane) or required assistance for activities of daily living prior to their stroke.

Withdrawal criteria. Subjects were withdrawn from the trial if they suffered any significant medical complication and/or were readmitted to an acute hospital for more than 1 week.

Recruitment

We approached all persons admitted to Hampstead Rehabilitation Centre, Adelaide, South Australia, for rehabilitation poststroke between March 2002 and October 2003 who met the inclusion and exclusion criteria within the first 2 days after admission and invited them to participate in the study.

Group Allocation

Within 3 days of admission to rehabilitation, we assigned consenting subjects to their treatment group. To ensure a sufficient number of circuit class participants at any one time for circuit class therapy sessions, it was not feasible to randomly allocate individual subjects to treatment groups. Instead, we allocated subjects in blocks according to the date of admission to rehabilitation. Subject recruitment occurred between March 2002 and October 2003 in 4 alternating blocks of time. Because of this, we were not able to conceal assignment to groups from the allocator, or blind the subjects or treating therapists to group allocation.

Evaluations

We assessed subjects on all outcome measures at admission, week 4 of rehabilitation, discharge from rehabilitation, and 6 months poststroke. An examiner who was unaware of the design and aim of the study and was blinded to subjects’ group allocation completed all subject assessments. The examiner remained blinded to the study design and aims, and subjects’ group allocation, until after data collection ceased. We considered the 4-week assessment as the primary end point of interest, as it was the average LOS for stroke rehabilitation at Hampstead Rehabilitation Centre at the time of the trial.

Measurements

Our primary outcome measures were gait speed measured by the five-meter walk test (5MWT), functional walking capacity as measured by the two-minute walk test (2MWT), and functional balance as measured by the Berg Balance Scale (BBS). For the 5MWT the assessor asked subjects to walk at a comfortable pace along a corridor which was marked at 3, 8, and 10m from the start position, without verbal encouragement. The assessor used a stopwatch to record the time taken to walk the middle 5m of the walkway. The 2MWT was conducted along the same walkway with large orange cones placed 1 and 9m from the start position. The assessor asked subjects to walk up and down the walkway, walking around the cones at each end, continuously for 2 minutes, at a comfortable pace, taking rests if required. The assessor counted the number of completed laps and at the end of the 2 minutes (measured on a stop watch) the subject was asked to stop and stand still. The walkway was marked at 0.5-m intervals and the assessor recorded the distance walked to the nearest 0.5m. In both the 5MWT and 2MWT, the assessor walked behind the subject at all times to avoid influencing the speed at which they walked. The BBS was assessed according to the standard protocol using the same equipment (stool, chairs, plinth) for each assessment.

Our secondary outcome measures were upper-limb function as measured by the upper-limb subscale of the Motor Assessment Scale (MAS) for stroke, which consists of the 3 upper-limb items added together to create 1 summary score, the degree of physical assistance required to walk as measured by the Iowa Level of Assistance Scale (ILAS), LOS in rehabilitation, and patient satisfaction as measured by a stroke-specific satisfaction questionnaire (adapted from Pound et al). We measured LOS as the number of days between admission and discharge from Hampstead Rehabilitation Centre.

All of the outcome measures have been shown to have acceptable interrater reliability and construct validity and were applied using standardized protocols. Intrarater reliability of the primary outcome measures was measured by the assessor re-scoring videotaped performances of the first 10 assessments 4 weeks later.

In addition, we obtained information regarding sociodemographic and stroke characteristics from the subjects’ medical records, with the level of stroke severity measured at admission to rehabilitation using the FIM instrument, administered by the treating multidisciplinary team.

Interventions

Subjects received the allocated type of PT for the duration of their inpatient stay and non-PT components of multidisciplinary rehabilitation were provided to both groups per usual practice. The broad aim of both individual and circuit class treatment sessions was to improve subjects’ mobility and upper-limb function to allow safe discharge to either their own home or appropriate supportive accommodation. The treating therapist recorded details regarding the content and duration of each therapy session. Subjects received their allocated PT intervention for the duration of their inpatient stay. Independent practice outside of individual therapy times was not usual practice at Hampstead Rehabilitation Centre at the time and therefore was not encouraged for either group during the trial.

We provided circuit class therapy to groups of up to 6 patients at any one time with a maximum of 1 physiotherapist and 1 PT assistant providing supervision in each class. If there were less than 5 participants in the class, only 1 physiotherapist provided supervision. The same physiotherapist (with 5 years of experience) provided supervision. The same physiotherapist (with 5 years of experience) provided supervision.
of experience in stroke rehabilitation) supervised all circuit classes, although several different PT assistants were involved. Circuit class participants attended two 90-minute treatment sessions a day, 5 days a week, and performed a set of core activities that addressed their key impairments and functional limitations. These core activities were: sit-to-stand practice; strengthening lower-limb extensor muscles in weight-bearing positions; postural control in standing; walking practice including negotiating obstacles, steps, ramps, stairs, and outdoor surfaces; reach and grasp; and fine manipulation of everyday household items in both unilateral and bilateral tasks. These core activities were individually adapted for each subject by the treating therapist and progressed as required, such that the level of difficulty, complexity, and dosage (number of repetitions) matched each individual’s ability. For example, practice of sit-to-stand ranged from starting with the subject perched on the edge of a high plinth with 1 therapist assisting, to independent practice using a low chair with no assistance while holding a glass of water. Appendix 1 includes a comprehensive list of exercises included in the class. We incorporated walking on a treadmill (without body-weight support) into the class for each subject as soon as they were able to walk overground with minimal assistance. We progressed treadmill training by increasing time (up to a maximum of 20 min per session) and speed. Group activities such as relay races were incorporated into the majority of classes. Although treating therapists provided some verbal feedback during circuit class therapy, the majority of feedback was provided by setting up each task such that it provided a concrete goal and intrinsic feedback regarding its correct completion.

Individual therapy sessions occurred under the direct and constant supervision of a physiotherapist or PT assistant, on a 1 therapist to 1 subject ratio, for up to 60 minutes a day, 5 days a week. Several different physiotherapists with a range of experience in stroke rehabilitation were involved in providing individual therapy. Individual therapy sessions were not based on any particular treatment philosophy and were tailored to the individual based on the physiotherapist’s assessment. Therapists often used manual guidance and verbal feedback for correct completion of tasks.

Sample Size

Prospective power calculations based on the ability to detect a between-group difference of 0.2 m/s in walking speed with 80% power (α = .05) indicated that a sample size of 37 subjects per group was required. We considered 0.2 m/s to be the minimum clinically important difference in walking speed, based on previous findings of an inherent measurement error of approximately 0.17 m/s with repeated measurements of walking on previous findings of an inherent measurement error of 0.17 m/s with repeated measurements of walking speed, thereby minimizing bias resulting from an analysis restricted to complete cases. When significant group by time interactions were found, we performed a post hoc pairwise comparison of mean scores to examine when these differences occurred. The post hoc tests used the same data set as that used for the linear mixed-model analyses—that is, it involved imputed values for missing data. We calculated 95% confidence intervals (CIs) for mean differences to further examine both significant and nonsignificant findings. Although the BBS is strictly an ordinal scale, we treated the data as continuous, because the range in scores was large and the distribution of scores closely approached normality. We used independent t tests to analyze between-group differences for hospital LOS, duration of therapy, and content of therapy. We used the Mann-Whitney U test to analyze between-group differences in MAS upper-limb subscale scores, and the chi-square statistic to analyze between-group differences in ILAS scores and responses to the patient satisfaction questionnaire. Analyses of the MAS upper-limb subscale scores, ILAS, and patient satisfaction questionnaire were based on cross-sectional data, not change scores.

We used the SPSS for all analyses, with a significance level of α equal to .05.

RESULTS

We conducted the trial between March 2002 and December 2003. Seventy-eight subjects consented to participate and were allocated to groups. Figure 1 depicts the progression of subjects through the trial and the reasons for exclusion and withdrawal. Ten subjects withdrew before the week-4 assessment and a further 4 subjects withdrew between week 4 and discharge. Table 1 presents reasons for the 14 subjects who were withdrawn from the study. One subject withdrew from the circuit class therapy group after week 4 due to safety concerns. Although this subject received individual therapy for the remainder of her inpatient stay, we included her discharge and follow-up data in the circuit class therapy group for all analyses according to intention to treat principles. Only 43 (63.2%) subjects attended the follow-up assessment. Those subjects who did attend had a faster walking speed at discharge from rehabilitation than those who did not attend (attendees, 0.81 ± 0.43 m/s; nonattendees, 0.59 ± 0.38 m/s; P = .04; t = −2.08, 95% CI, 0.01–0.46). There were no significant differences in discharge 2MWT scores and BBS data between those subjects who did and did not attend the follow-up assessment.

Intrater reliability for the primary outcome measures was high, with intraclass correlation coefficient scores of 1.00, 1.00, and 0.96 for the 5MWT, 2MWT, and BBS scores, respectively.

Table 2 presents baseline characteristics of subjects included in data analyses. Data from the 5MWT and 2MWT were highly skewed at admission due to 21 (30.9%) subjects being unable to walk. There were no significant differences between treatment groups at baseline for any of the primary outcome measures or the admission FIM score. However, circuit class therapy subjects were significantly younger (individual therapy, 68.9 ± 12.3y; circuit class therapy, 61.6 ± 11.8y; P = .02).

Table 3 presents summary data for all outcome measures. The linear mixed-model analyses demonstrated a significant group by time interaction effect for each of the 3 primary outcome measures (5MWT: F = 2.88, P = .04; 2MWT: F = 6.82, P < .001; BBS scores: F = 2.79, P = .04) suggesting that the
treatment groups behaved differently over time. Post hoc pair-wise comparisons of mean scores demonstrated that both groups showed significant improvement between admission and week 4 on the 5MWT (mean improvement: individual therapy, .16m/s; circuit class therapy, .17m/s), the 2MWT (mean improvement: individual therapy, 21.3m; circuit class therapy, 16.5m), and the BBS (mean improvement: individual therapy, 8.9 points; circuit class therapy, 8.5 points). Both groups continued to show significant improvement in all 3 measures between week 4 and discharge; however, only subjects receiving individual therapy showed significant improvement in the 5MWT and 2MWT between discharge and follow-up. There were no significant between-group differences at any time point, with the exception of the follow-up assessment at which subjects in the individual therapy group scored significantly higher on the 2MWT. Analyses were repeated with age as a covariate, which did not alter results.

There were no significant differences between groups on the MAS upper-limb subscale at any time point. There was a low frequency of scores on the ILAS for the majority of categories; therefore, data were collapsed into 2 categories (independent, assistance required/unsafe to attempt assessment) prior to analysis. At discharge a significantly higher proportion of subjects in the circuit class therapy group (n = 26 [92.9%]) were able to walk independently compared with the individual therapy group (n = 23 [67.6%], $\chi^2$ test = 5.89, $P = .01$). There were no significant differences between groups in the responses to the patient satisfaction questionnaire at discharge or follow-up.

Fig 1. Flow chart of trial participants. *These subjects were not withdrawn, but did not complete either the week-4 or discharge assessment. †Includes 1 subject in each treatment group with LOS of less than 4 weeks, therefore did not complete a week-4 assessment. ‡One subject refused a discharge assessment, but completed a follow-up assessment.
with 1 exception. At follow-up, significantly more subjects in the circuit class therapy group agreed or strongly agreed (n=21 [95.5%]) that they had received enough PT during their inpatient rehabilitation stay compared with the individual therapy group (n=11 [55.0%], χ² test=12.27, P=.003) (see Table 3).

We calculated LOS for all 64 subjects who completed the trial, including 2 subjects who refused to complete the discharge assessment. Discharge was delayed (due to waiting either for a bed in a residential care facility or for essential home modifications to be completed) for 8 subjects (5 in the individual therapy group, 3 in the circuit class therapy group). LOS for these subjects was calculated as the number of days between admission to rehabilitation and the date of the discharge assessment (this assessment occurred within 1 week of the multidisciplinary team deciding that the patient was ready for discharge). Although subjects in the circuit class therapy group had a shorter mean LOS (mean difference, 15.2d), this did not reach statistical significance (95% CI, −4.5 to 34.7d).

Subjects in the circuit class therapy group received significantly more therapy time per day (mean, 129.1 ± 22.6 min) than subjects in the individual therapy group (mean, 36.6 ± 9.4 min) (P<.001, t=−22.68; 95% CI, −101.1 to −83.7). Subjects in the individual therapy group received an average of 49.0 therapy sessions with a mean of 0.94 sessions per day. Subjects in the circuit class therapy group received an average of 60.0 therapy sessions with a mean of 1.6 sessions per day. On average there were 4 subjects in the circuit class at any one time. Both types of therapy were well tolerated with an attendance rate of 92.1% in the individual therapy group and 83.3% in the circuit class therapy group. Two subjects fell during individual therapy sessions and 4 subjects fell during circuit class therapy sessions. None of these falls caused injury.

We analyzed data from the first 20 treatment sessions for each subject in order to accurately describe the content of individual therapy and circuit class therapy, and to determine key differences in therapy content between the groups (Table 4). More than 75% of circuit class sessions included walking and sit-to-stand practice and approximately a third of all circuit class sessions included treadmill training. By contrast, only a third of individual sessions included walking practice and treadmill training was never included. However, almost half of all individual treatment sessions included activities directed at improving the quality of a subject’s gait (eg, part-practice of the swing phase of gait with a focus on reducing compensatory hip circumduction or part-practice of the stance phase of gait with a focus on increasing weight shift to the paretic side), whereas gait quality was not specifically addressed in any circuit class therapy sessions. Practice of upper-limb functional activities—such as lifting a cup to the mouth—was included in a quarter of both circuit class therapy and individual therapy sessions. In terms of the key statistical differences in therapy content between groups, significantly more circuit class therapy sessions included practice of sit-to-stand (P<.001, t=−6.32; 95% CI, 12.3 to 6.4), walking (including negotiating obstacles and stairs, P<.001, t=−7.01; 95% CI, −11.5 to −6.4) and treadmill training (P<.001, t=−5.78; 95% CI, −7.7 to −4.0). Significantly more individual therapy sessions included practice of transfers (including bed-to-chair, bed-to-toilet, and car transfers, but not including sit-to-stand practice, P<.001, t=4.37; 95% CI, 1.7–5.2) and exercises aimed at improving gait quality (P<.001, t=7.68; 95% CI, 6.3–11.1).

**DISCUSSION**

We found no clinically important or statistically significant differences between circuit class therapy and individual PT sessions for recovery of walking ability, functional balance, and upper-limb function among a sample of persons receiving inpatient rehabilitation poststroke. However, a significantly greater number of subjects in the circuit class therapy group were able to walk independently at discharge from rehabilitation and were satisfied with the amount of therapy received compared with subjects who received individual therapy. The sample included in our study can be considered representative of a larger population of persons receiving inpatient rehabilitation after stroke because it included a broad range of functional disability at baseline.

Despite the fact that subjects in the circuit class therapy group did not receive individual treatment sessions, there were few significant differences between the groups in terms of walking ability, functional balance, and upper-limb function, with subjects in both groups showing significant improvement over time. In addition, circuit class therapy was more effective than individual therapy in promoting independence in walking at discharge from rehabilitation in this sample of subjects. This may have been due to circuit class therapy subjects spending a significantly greater amount of time engaged in walking practice. Additionally, the structure of circuit class therapy was such that it encouraged greater participant autonomy, thereby encouraging problem solving and independence. Thus, it is possible that not only the increased amount of walking practice, but also the environment in which this practice occurred, had an impact on the increased level of walking independence.

### Table 1: Reasons for Subject Attrition

<table>
<thead>
<tr>
<th>Reason for Attrition</th>
<th>Individual Therapy</th>
<th>Circuit Class Therapy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suspected lower-limb fracture after fall on ward</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Suspected extension of stroke</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Withdrawal of informed consent</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Lower-limb vascular surgery required</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Acute illness requiring readmission to an acute hospital (decreased respiratory function, severe urinary tract infection, acute bowel obstruction)</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Discharged within 2 weeks</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Acute psychiatric illness requiring readmission to an acute hospital</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Total number of subjects lost</td>
<td>4</td>
<td>10</td>
</tr>
</tbody>
</table>

**NOTE.** Values are counts. Ten subjects withdrew before the week-4 assessment and a further 4 subjects withdrew between week 4 and discharge.

### Table 2: Baseline Characteristics of the Study Sample

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Individual Therapy (n=37)</th>
<th>Circuit Class Therapy (n=31)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age ± SD (y)</td>
<td>68.9±12.3</td>
<td>61.6±11.8*</td>
</tr>
<tr>
<td>Mean FIM score ± SD (range, 18–126)</td>
<td>78.7±17.6</td>
<td>83.1±16.5</td>
</tr>
<tr>
<td>Sex (male/female)</td>
<td>25/12</td>
<td>16/15</td>
</tr>
<tr>
<td>Side of stroke lesion (left/right)</td>
<td>15/22</td>
<td>11/20</td>
</tr>
<tr>
<td>Stroke type (infarct/hemorrhage)</td>
<td>34/3</td>
<td>24/7</td>
</tr>
<tr>
<td>Mean time between stroke and admission to rehabilitation ± SD (d)</td>
<td>24.4±12.4</td>
<td>29.7±15.5</td>
</tr>
</tbody>
</table>

*Statistically significant difference at P<.05.
within the circuit class therapy group. All known previous studies of circuit class therapy involved subjects who were already able to walk independently at baseline.\(^\text{14,16,38-40}\) Therefore, this study is the first to show that circuit class therapy is effective in promoting independent walking ability poststroke. The reasons for the significant improvement in gait speed and functional walking capacity between discharge and follow-up for the individual therapy group but not the circuit class therapy

### Table 3: Summary of Main Outcome Variables

<table>
<thead>
<tr>
<th>Outcome Variables</th>
<th>Individual Therapy</th>
<th>Circuit Class Therapy</th>
<th>Mean Difference Between Groups (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5MWT (m/s)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean ± SD (n)</td>
<td>0.37 ± 0.40 (37)</td>
<td>0.41 ± 0.43 (31)</td>
<td>−0.04 (−0.25 to 0.17)</td>
</tr>
<tr>
<td>Admission</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 4</td>
<td>0.53 ± 0.43 (36)</td>
<td>0.58 ± 0.46 (30)</td>
<td>−0.07 (−0.28 to 0.14)</td>
</tr>
<tr>
<td>Discharge</td>
<td>0.72 ± 0.43 (34)</td>
<td>0.76 ± 0.41 (28)</td>
<td>−0.02 (−0.20 to 0.23)</td>
</tr>
<tr>
<td>Follow-up</td>
<td>0.95 ± 0.45 (20)</td>
<td>0.85 ± 0.48 (22)</td>
<td>0.13 (−0.09 to 0.35)</td>
</tr>
<tr>
<td>2MWT (m)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean ± SD (n)</td>
<td>36.8 ± 40.3 (37)</td>
<td>41.1 ± 44.1 (31)</td>
<td>−4.3 (−26.3 to 17.6)</td>
</tr>
<tr>
<td>Admission</td>
<td>58.1 ± 48.7 (36)</td>
<td>57.6 ± 44.9 (30)</td>
<td>−1.8 (−23.8 to 20.2)</td>
</tr>
<tr>
<td>Week 4</td>
<td>74.3 ± 46.5 (34)</td>
<td>76.1 ± 43.0 (28)</td>
<td>3.1 (−18.9 to 25.2)</td>
</tr>
<tr>
<td>Discharge</td>
<td>103.0 ± 47.6 (20)</td>
<td>82.1 ± 48.3 (22)</td>
<td>23.1 (0.21 to 46.0)*</td>
</tr>
<tr>
<td>Follow-up</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BBS scores (range, 0–56)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean ± SD (n)</td>
<td>28.2 ± 17.7 (37)</td>
<td>32.3 ± 15.5 (31)</td>
<td>−4.1 (−10.5 to 2.3)</td>
</tr>
<tr>
<td>Admission</td>
<td>37.1 ± 16.4 (36)</td>
<td>40.8 ± 12.9 (30)</td>
<td>−3.9 (−10.3 to 2.0)</td>
</tr>
<tr>
<td>Week 4</td>
<td>46.7 ± 7.9 (34)</td>
<td>48.0 ± 7.4 (28)</td>
<td>1.1 (−5.2 to 7.8)</td>
</tr>
<tr>
<td>Discharge</td>
<td>50.8 ± 4.9 (20)</td>
<td>49.1 ± 8.5 (22)</td>
<td>3.0 (−4.1 to 10.1)</td>
</tr>
<tr>
<td>Follow-up</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upper-limb MAS scores (range, 0–18)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean ± SD (n)</td>
<td>5.0 (0.8–12.0) (37)</td>
<td>10.0 (1.0–13.0) (31)</td>
<td>NA</td>
</tr>
<tr>
<td>Admission</td>
<td>6.5 (1.0–13.0) (36)</td>
<td>12.0 (1.0–14.0) (30)</td>
<td>NA</td>
</tr>
<tr>
<td>Week 4</td>
<td>9.5 (4.8–14.0) (34)</td>
<td>12.0 (2.0–14.0) (28)</td>
<td>NA</td>
</tr>
<tr>
<td>Discharge</td>
<td>13.0 (7.5–14.8) (20)</td>
<td>14.0 (2.0–16.0) (22)</td>
<td>NA</td>
</tr>
<tr>
<td>Follow-up</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of subjects rated as independent on the ILAS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N (%) (n)</td>
<td>7 (18.9) (37)</td>
<td>7 (22.6) (31)</td>
<td>NA</td>
</tr>
<tr>
<td>Admission</td>
<td>14 (38.9) (36)</td>
<td>15 (50.0) (30)</td>
<td>NA</td>
</tr>
<tr>
<td>Week 4</td>
<td>23 (67.6) (34)</td>
<td>26 (92.0) (28)</td>
<td>NA</td>
</tr>
<tr>
<td>Discharge</td>
<td>17 (85.0) (20)</td>
<td>19 (61.3) (22)</td>
<td>NA</td>
</tr>
<tr>
<td>Follow-up</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Responses to the question “I have had enough physiotherapy” at follow-up</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N (%)</td>
<td>10 (50.0)</td>
<td>13 (59.1)</td>
<td>NA</td>
</tr>
<tr>
<td>Strongly agree</td>
<td>1 (5.0)</td>
<td>8 (36.4)</td>
<td>NA</td>
</tr>
<tr>
<td>Agree</td>
<td>7 (35.0)</td>
<td>1 (4.5)</td>
<td>NA</td>
</tr>
<tr>
<td>Disagree</td>
<td>2 (10.0)</td>
<td>0 (0.0)</td>
<td>NA</td>
</tr>
<tr>
<td>Strongly disagree</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOS (d)</td>
<td>71.3 ± 44.0 (36)</td>
<td>56.1 ± 31.1 (28)</td>
<td>15.2 (−4.5 to 34.7)</td>
</tr>
</tbody>
</table>

Abbreviations: IRQ, interquartile range; NA, not applicable.
*Statistically significant between-group difference at \(P \leq 0.05\).

### Table 4: Summary of the Activities of Individual and Circuit Class Subjects Over 20 Treatment Sessions Showing Means and Between-Groups Differences

<table>
<thead>
<tr>
<th>Activities</th>
<th>Individual Therapy</th>
<th>Circuit Class Therapy</th>
<th>Mean Difference (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treadmill training</td>
<td>0 ± 0</td>
<td>5.84 ± 5.62</td>
<td>−5.8 (−7.7 to −4.0)*</td>
</tr>
<tr>
<td>Gait quality</td>
<td>9.14 ± 6.49</td>
<td>0.42 ± 2.16</td>
<td>8.7 (6.3 to 11.1)*</td>
</tr>
<tr>
<td>Gait practice</td>
<td>6.27 ± 5.36</td>
<td>15.19 ± 5.06</td>
<td>−8.9 (−11.5 to −6.4)*</td>
</tr>
<tr>
<td>Upper-limb functional activities</td>
<td>4.97 ± 5.24</td>
<td>4.81 ± 6.30</td>
<td>−0.17 (−2.7 to 3.0)</td>
</tr>
<tr>
<td>Sit to stand</td>
<td>5.70 ± 5.62</td>
<td>15.06 ± 6.60</td>
<td>−9.4 (−12.3 to −6.4)*</td>
</tr>
<tr>
<td>Transfer practice†</td>
<td>3.78 ± 4.74</td>
<td>0.32 ± 0.79</td>
<td>3.5 (1.7 to 5.2)*</td>
</tr>
</tbody>
</table>

*Statistically significant difference at \(P \leq 0.05\).
†Includes bed-to-chair, bed-to-toilet, and car transfers, but not sit-to-stand.
CIRCUIT CLASS THERAPY FOR STROKE REHABILITATION, English

The ability to provide a significantly greater amount of therapy time with a lower staff to patient ratio in our study suggests that circuit class therapy may also be a more cost-effective method of therapy delivery. In our study the total amount of therapist time required to provide circuit class therapy for 6 patients was a mean of 130 minutes a day, whereas the total amount of therapist time required to provide individual therapy sessions for 6 patients was 222 minutes a day (based on a mean duration of 37 minutes for individual sessions). This represents a difference of more than 90 minutes a day of therapist time. Additionally, circuit class therapy participants received a significantly greater amount of therapy a day. This additional therapy time may have resulted in a shorter hospital LOS, because although the difference in LOS in our study was not statistically significant, the confidence interval suggested a trend in favor of circuit class group subjects (mean difference, 15.2d; 95% CI, −4.5 to 34.7d). Coupled with similar findings by Blennerhassett and Dite of a large, but not statistically significant, reduction in LOS associated with mobility related to circuit class therapy, it is clear that future research into the cost-effectiveness of circuit class therapy as an alternative method of service provision is required.

There is increasing emphasis being placed on the importance of patient satisfaction within health care evaluation and evidence based medicine. Several studies have found that the majority of persons recovering from stroke were not satisfied with the medical and rehabilitative care they received. Similarly, in our study 45% of subjects receiving individual therapy were not satisfied with the amount of therapy they received. Our finding that 95.5% of subjects receiving circuit class therapy were satisfied with the amount of PT they received is of particular clinical relevance, although perhaps not surprising, considering the increased amount of therapy time they received. However, this finding must be interpreted with caution in view of the large drop-out rate between discharge and follow-up. Nevertheless, patient satisfaction data collected at follow-up may be a more accurate reflection than that measured during inpatient rehabilitation, given that patients are often reluctant to express dissatisfaction with hospital services, and may therefore be more honest in reporting levels of satisfaction when they are no longer receiving hospital services.

Study Limitations

The heterogeneous nature of the sample was a limitation of our study. Further research with a larger, more homogeneous sample of subjects with moderate disability may yield more definitive results, as it is well established that persons with stroke resulting in moderate disability demonstrate greater improvement in functional abilities during rehabilitation, compared with those with mild or severe impairment. Although the nonrandom allocation of subjects to groups in our study had the potential to bias the results, age was the only significant difference between the treatment groups at baseline, and the use of age as a covariant in the analyses did not alter the results. Although an additional limitation was the rate of subject attrition, the reasons for withdrawal from the study (see table 1) appeared to be unrelated to the type of therapy provided. Although 3 subjects suffered falls necessitating withdrawal from the study, all occurred outside of therapy sessions, making it highly improbable that the type of PT the subjects received was responsible. Additionally, although there were 2 cases of suspected extension of stroke, both occurring in the circuit class therapy group, neither was confirmed by radiologic examination. Some studies in the rat model have suggested that early, intensive therapy within the first week after stroke may increase the risk of stroke symptoms worsening. However, there have been no published reports of stroke extension or worsening stroke symptoms associated with increased intensity of therapy provided to persons more than a week after stroke. Nevertheless, it would be pertinent for future studies of circuit class therapy to include assessment of the incidence and severity of infarct extension. A further limitation of our study design was that the discharge assessment did not occur at a standardized time. However, the inclusion of these data made it possible to compare functional ability at discharge between treatment groups. These limitations of study design highlight the difficulties inherent in clinical research, in particular finding a balance between ideal study design, the practicalities of clinical research, and the applicability of the findings to a clinical setting.

Future research into the effectiveness of circuit class therapy is clearly warranted. Preferably, such research should involve a more homogeneous study sample, random allocation to treatment groups, and standardized assessment times. Although such studies would require significant financial support, the potential for circuit class therapy to be a more cost effective method of service delivery and to reduce hospital LOS justifies this support.

CONCLUSIONS

Subjects receiving either circuit class therapy or individual therapy demonstrated a similar degree of recovery on objective measures of mobility and upper-limb function for persons receiving inpatient rehabilitation after stroke. However, circuit class therapy was associated with a significantly greater degree of independence in walking at discharge from rehabilitation and significantly higher patient satisfaction with the amount of therapy received. Furthermore, the study demonstrated the feasibility and safety of circuit class therapy as an alternative sole model of PT service delivery for person receiving inpatient rehabilitation after stroke.

Acknowledgments: We thank the staff of the Royal Adelaide Hospital Physiotherapy Department and Medical Rehabilitation Unit at Hampstead Rehabilitation Centre, in particular Julie McGuiness, for their support and assistance in conducting the trial.

APPENDIX 1: LIST OF EXERCISES INCLUDED IN CIRCUIT CLASS THERAPY

Lower-limb exercises
- Forward and lateral step raises. (Subject placed affected leg on a step placed in front or to the side and raised him/herself onto step.)
- Eccentric quads over edge of step. (Subject stood on a step and lowered unaffected leg to touch the ground. To progress exercise, subjects touched a foam cup placed on
APPENDIX 1: LIST OF EXERCISES INCLUDED IN CIRCUIT CLASS THERAPY (cont’d)

the ground in front of the step without crushing it before returning to the start position.)
• Heel raises, either from standing flat on the ground or standing on a wedge.
• Squats.
• Active hamstrings in sitting. (Subject sat on a chair and bent affected knee as far as possible. A small towel or “slippery sam” material was used to reduce friction to make the exercise easier and weights were strapped to the ankle to make the exercise harder.)
• Reaching in various directions out of the base of support in sitting (to activate leg musculature).
• Stretches of the gastrocnemius and soleus muscles in standing.
• Riding an exercise bicycle.

Sit-to-stand and walking exercises
• Sit-to-stand from various heights including from an adjustable plinth, chairs with and without arms, and stools. To progress the exercise, subjects placed their unaffected leg on a small step.
• Stepping onto and off a step.
• Walking indoors, forward, backward, and sideways.
• Walking outdoors.
• Walking on the treadmill.
• Walking up and down stairs.
• Walking around obstacle courses including stepping over and around objects, up and down steps, over soft surfaces, and picking up objects from the floor. Subjects also negotiated obstacle courses while carrying a tray of objects (for dual-task performance).

Exercises to improve postural control in standing
• Reaching to touch marks or trace a spiral shape on a whiteboard. Different stance positions were used including feet together and tandem stance.
• Stepping grid. Subjects stood with feet in marked areas, then tapped 1 foot out to touch marks on floor, repeating with the other foot.
• Reaching to pick up objects from low surfaces or the floor. This exercise was also performed in pairs with subjects passing objects to each other.
• Alternatively tapping toes onto step in front. To progress the exercise, subjects had to tap a foam cup placed on the step without deforming it.
• Throwing and catching balls in pairs or groups.
• Practicing standing on 1 leg.
• Walking with a heel-toe gait or “braiding” (ie, walking sideways with the trailing leg alternating between crossing in front of, or behind the leading leg with each step).

Exercises for the upper limb and hand
• Prolonged shoulder positioning in either forward flexion or abduction. Where elbow contracture or stiffness was an issue, a circumferential foam splint was used to hold the elbow in maximal extension.
• Active shoulder girdle movement with the arm supported on a high surface.
• Reaching to grasp and move various size objects on and off of “shelves.”
• Taking lids on and off jars.
• Pegging washing onto a line.
• Folding washing.

References
16. Pang M, Eng J, Dawson A, McKay H, Harris J. A community-based fitness and mobility exercise program for older adults with...


37. Little R, Raghunathan T. On summary measures analysis of the linear mixed effects model for repeated measures when data are not missing completely at random. Stats Med 1999;18:2465-78.


Kinematic and Clinical Analyses of Upper-Extremity Movements After Constraint-Induced Movement Therapy in Patients With Stroke: A Randomized Controlled Trial

Ching-yi Wu, ScD, OTR, Chia-ling Chen, MD, PhD, Simon F. Tang, MD, Keh-chung Lin, ScD, OTR, Ya-ying Huang, MS


Objective: To study the effects of constraint-induced movement therapy (CIMT) relative to traditional intervention on motor-control strategies for upper-arm reaching and motor performance at the impairment and functional levels in stroke patients.

Design: Two-group randomized controlled trial (RCT): pretreatment and posttreatment measures.

Setting: Rehabilitation clinics.

Participants: Forty-seven stroke patients (mean age, 55y) 3 weeks to 37 months postonset of a first-ever cerebrovascular accident.

Interventions: Forty-seven patients received either CIMT (restraint of the less affected hand combined with intensive training of the more affected upper extremity) or traditional intervention (control treatment) during the study. The treatment intensity was matched between the 2 groups (2h/d, 5d/wk for 3wk).

Main Outcome Measures: Outcomes were evaluated using (1) kinematic variables of reaching movement used to describe the control strategies for reaching, (2) the Fugl-Meyer Assessment (FMA) of motor-impairment severity, and (3) the Motor Activity Log (MAL) evaluating the functional ability of the upper extremity.

Results: After treatment, the CIMT group showed better strategies of reaching control than the control group (P <.03). The CIMT group also showed less motor impairment on the FMA (P = .019) and higher functional ability on the MAL (P <.001).

Conclusions: This is the first RCT to show differences in motor-control strategies as measured by kinematic variables after CIMT versus traditional intervention. In addition to improving motor performance at the impairment and functional levels, CIMT conferred therapeutic benefits on control strategies determined by kinematic analysis.

Key Words: Controlled clinical trials; Kinematics; Occupational therapy; Rehabilitation; Stroke.

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

With the proliferation of successful controlled clinical trials of constraint-induced movement therapy (CIMT) in the literature, the therapy has attracted considerable attention as a means to treat the more affected upper extremity (UE) and overcome the learned nonuse phenomenon (habitually relying on the less affected UE to accomplish functional tasks) among patients with stroke.1–5 CIMT involves restraint of the less affected UE over an extended period (6–18h/d for 2–3wk), in combination with intensive task-specific training of the more affected limb (eg, 6h/d on 10–15 consecutive weekdays).6 In contrast to clinical efficacy of CIMT, the acceptance of CIMT among therapists and patients remains poor. One possible reason is that intense and massed practice during CIMT may be less safe and more tiring for stroke patients.6 To address the problems, a variety of derivatives of CIMT were devised by using less intensive training and shorter restraint time. These derivatives of CIMT may involve training of the affected UE for 0.5 to 3 hours a day, 3 to 5 days a week for 2 to 10 weeks, together with restraint of the less affected UE for up to 6 hours a day.7–14 Substantial evidence shows that CIMT can improve motor impairment and functional use of the more affected UE in chronic (>1y poststroke),8,9 subacute (>3mo <1y poststroke),7,10,11 and acute10,12,14 stroke patients with mild motor impairments. However, the motor-control mechanisms responsible for the improved motor performance are poorly understood. A possible reason for this lack of understanding relates to the methods used to assess UE function.15 Most studies of CIMT have relied on clinical evaluations such as observer-initiated measure of motor impairment (eg, the Fugl-Meyer Assessment [FMA]) or self-report measure of functional ability (eg, Motor Activity Log [MAL]).7–9,11–14 These tests provide clinical information regarding the level of motor impairment and the level of functional ability or patients’ perceptions of how well they can functionally use their UEs. However, these measures may have particular aspects that are left up to the subjective perspective of the rater.12,15 Technologic developments in motion analysis enabled biomechanics studies of actions (eg, upper-arm reaching). Biomechanic analysis of movements provides more objective and quantitative measures of control strategies necessary for promoting understanding of the mechanisms underlying improved motor performance of
stroke patients after CIMT. Recent research has recommended the use of biomechanic analysis for outcome evaluations of CIMT.12

Alberts et al13 used kinetic analysis to objectively examine how CIMT affects control of force in precision grip. In addition to kinetic analysis, kinematic analysis during functional tasks has been suggested as a valid means to directly and objectively measure the spatiotemporal control of movement.14 Kinematic measures of reaching have been shown to discriminate changes to hand-path quality after brain injury15 and are associated with measures of arm functions (eg, the Action Research Arm Test).16 Such measures can be used not only to assess performance but also to identify movement quality or to elucidate the motor strategies during a reaching task.17 The information may afford insights into how motor-control rehabilitation alters movement organization in patients with stroke.

A normal-reaching movement is controlled by both preprogramming and online error correction.20-22 A more preprogrammed control strategy for reaching movement represents a better learned or more skillful movement. When the movement depends more on motor preprogramming, it will be more rapidly initiated, more efficient and direct, and smoother. Preprogrammed movement requires rapid task completion, such that the performer may need greater force or impulse at movement initiation to quickly bring the hand to the target. In contrast, if the control strategy of the movement depends more on online correction, it will require more planning and therefore be initiated more slowly. This type of movement requires adjustment of direction based on sensory feedback during performance, resulting in reduced movement efficiency, straightness, and smoothness.

A recent case report23 has used kinematic measures to evaluate the changes in motor control after CIMT and presented preliminary evidence that more preprogrammed control strategy of reaching movement is achieved after than before CIMT. Further research performed by using a randomized controlled trial (RCT) is needed to afford insights into mechanisms mediating change in motor performance after CIMT. The present study used kinematic analysis to investigate the impact of CIMT versus traditional intervention on control strategies for reaching movement and on motor impairment and functional ability of the UE after stroke. Combining kinematic analysis and clinical evaluation might enable comprehensive assessment of the change in control strategies and motor performance at the impairment and functional levels after CIMT. Reaching was selected as the study task because of its status as a fundamental motor skill necessary for many daily activities. We hypothesized that patients receiving 3 weeks of CIMT would exhibit better strategies of reaching control, reduced motor impairment, and increased functional ability involving the more affected UE than patients receiving the traditional intervention. Better control strategies would be evidenced by more efficient, straighter, and smoother movements. Reduced motor impairment and increased functional ability would be represented by higher scores of FMA and MAL.

METHODS

Participants

The institutional review boards for human studies at 2 medical centers approved this protocol, and all subjects gave informed consent. Subjects were recruited from 2 stroke rehabilitation units; 47 stroke patients (32 men, 15 women; mean age, 55y; range, 40–80y) participated. According to self-report, subjects were right-hand dominant before stroke. At the beginning of the intervention, they were 3 weeks to 37 months (mean, 12.25mo) postonset of a first-time cerebrovascular accident of ischemic or hemorrhagic type. Inclusion criteria were as follows: (1) able to reach Brunnstrom stage24 III or above for the proximal part of the UE, (2) considerable nonuse of the more affected UE (MAL amount of use [AOU]25 score <2.5), (3) no serious cognitive deficits (modified Mini-Mental State Examination26 score ≥70), (4) no balance problems sufficient to compromise safety when wearing the experimental constraint device, and (5) lack of participation in any experimental rehabilitation or drug studies and absence of use of antispasticity drugs for UE musculature (eg, botulinum toxin type A) within the past 3 months. All patients received independent examinations by a physiatrist and an occupational therapist to determine their eligibility for inclusion. Figure 1 details subject recruitment and assignment.

Instruments

Movement data were collected using kinematic analysis and clinical evaluation as described later.

**Kinematic analysis.** A desk bell (diameter, 9.7cm [3.80in]; height, 4.8cm [1.87in]) was used as the target for reaching to press it. During reaching, a 6-camera motion analysis system4 was used in conjunction with a personal computer to capture the movement of a marker attached to the styloid process of the ulna; 2 channels of analog signals were collected simultaneously. The analog signals were connected to a pressure-sensitive hand switch and the desk bell. Movement recording began when the hand moved off the hand switch, and movement termination was recorded when the subject pressed the desk bell. Movements were recorded at 60Hz and digitally low-pass filtered at 5Hz using a second-order Butterworth filter with a forward and backward pass.

**Clinical evaluation.** The clinical measures used in this study included the FMA and the MAL. The FMA was used to assess several dimensions of motor impairment. The MAL is a...
functional measure and used to evaluate the ability of performing essential tasks and functional activities through self-report.

Scoring on the arm and hand section of the FMA (maximum score, 66) was based on a 3-point ordinal scale (0, cannot perform; 1, can perform partially; 2, can perform fully). A higher FMA score indicates less motor impairment. Test-retest reliability, interrater reliability, and construct validity of the test are well established.

The MAL is a semistructured interview that obtained information about how patients use their more affected limbs during 30 important daily activities. Patients were instructed to use 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 more affected UEs during the past week.

**Intervention**

Treatment regimens were designed to ensure that patients received equal treatment intensity (2h/d, 5d/wk for 3 consecutive weeks) directly supervised by the occupational therapists. The intervention was provided at 2 centers under the supervision of 2 separate occupational therapists. These 2 therapists were trained in the administration of the CIMT protocol by the investigators and completed a written competency test before subject treatment. During the treatment period, structured daily treatment notes were made and reviewed by the investigators to ensure the standardization of treatment. All trainings were provided on an individual basis. Subjects were blind to the study hypotheses. All subjects received routine interdisciplinary stroke rehabilitation separate from the study treatment that occurred during the regularly scheduled occupational therapy sessions. The interdisciplinary stroke rehabilitation was delivered by a variety of treatment disciplines including physical therapists and psychologists. The intensity of the interdisciplinary rehabilitation was the same for all participants (1.5h/d for 5d/wk).

In the CIMT group, typical training activities involved the use of the more affected UE and were similar to those performed daily (eg, reaching forward to move a jar from 1 place to another, picking up a cup and drinking from it, picking up a hairbrush and combing hair, cleaning the window). The less affected hand was placed in a mitt for 6 hours a day throughout the study period. In the traditional intervention group, the treatment involved neurodevelopmental therapy emphasizing functional task practice when possible, stretching and weight bearing with the more affected arm, and fine-motor dexterity training.

**Procedures and Testing**

Subjects were randomly assigned to the CIMT or traditional intervention group by using a random numbers table. Before and after the 3-week intervention period, the laboratory test (ie, kinematic analysis) and the clinical evaluation (FMA, MAL) were administered in random order by a blinded rater. The order of the laboratory test and the clinical evaluation was randomized to wash out carryover effects. Before administration of the outcome measures, a blinded rater was trained to administer the FMA and MAL following the guidelines described by Fugl-Meyer and Taub and colleagues, respectively. Rater competence was assessed by the primary investigators who had 5 years of experience in using these measures. The rater was trained to conduct the kinematic analysis in accord with standardized procedures described as follows.

During the laboratory test for reaching kinematics, each subject sat on a height-adjustable chair with seat-height set to 100% of the lower leg length, measured from the lateral knee joint to the floor with the subject standing. The trunk was secured to the chair back with a harness to prevent lateral and forward flexion and rotation but still allow for scapular motion. The subject rested the more affected hand on the hand switch placed on the edge of the table in line with the subject’s midsagittal plane. Table height was adjusted to 5.1cm (2 in) below the elbow. The desk bell was located along the subject’s midsagittal plane, and the reaching distance to the bell was standardized to the subject’s functional arm length. Functional arm length was defined as the distance from the medial border of the axilla to the distal wrist crease when the subject raised his/her arm as close to 90° elevation as possible and reached forward (without trunk movement) as far as possible. The functional arm length represents the farthest distance the subject can reach forward without using trunk movement. Subjects were instructed to reach and press the desk bell using the hand at a comfortable speed; 5 trials were performed after a practice trial.

**Data Reduction and Data Analysis**

An analysis program coded by LabView® language was used to process the kinematic data. Values of reaction time, movement time, total displacement, peak velocity, and movement units were obtained. Because the task distance varied across subjects, movement time, total displacement, and movement units were normalized to correct for variations in reaching distance.

Reaction time refers to the time to initiate the movement. Movement time means the time for execution of the reaching movement, representing movement efficiency. Total displacement refers to the path of the hand in 3-dimensional space, indicating the directness of movement. One movement unit is comprised of 1 acceleration and 1 deceleration phase and can be used to characterize movement smoothness and evaluate the extent of error correction during movement performance. Fewer movement units indicate smoother movement. Peak velocity indicates force or impulse at movement initiation. Greater force or impulse is reflected in a higher-amplitude peak velocity.

Analysis of covariance (ANCOVA), controlling for pretreatment differences, was used to test whether the CIMT group performed significantly better than the traditional intervention group on the posttest for each variable. For each analysis, pretest performance (kinematic data, FMA and MAL scores) and time postonset of stroke were used as the covariates, group was the independent variable, and posttest performance was the dependent variable. The effect size $r$ was calculated for each outcome variable to index the magnitude of the performance difference between groups. 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.

**RESULTS**

There were no significant differences between the 2 groups with respect to the measured characteristics except for the pretreatment kinematic performance on normalized total displacement (table 1). We addressed the pretest variability between groups by treating the pretreatment performance as a covariate in the comparison of posttest performance by the 2 groups.

The ANCOVA results showed significant and moderate effects on most kinematic measures at posttest (table 2). Patients treated with CIMT were able to initiate movement more quickly than patients treated with traditional intervention, as...
The CIMT group performed the task more efficiently (shorter normalized movement time, \( P = .005 \)) and with straighter (smaller normalized total displacement, \( P = .021 \)) and smoother (fewer normalized movement units, \( P = .010 \)) reaching trajectories of the more affected arm. Thus, the CIMT group showed more preprogrammed movement after treatment than did the traditional intervention group. A nonsignificant and small effect was found for the kinematic variable peak velocity.

The results showed significant and moderate to large effects in favor of the CIMT group on the FMA and MAL. There was less motor impairment for the CIMT group versus the traditional intervention group (FMA, \( P = .019 \)). Subjects in the CIMT group reported better performance in AOU (\( P < .001 \)) and QOM (\( P < .001 \)) of their more affected UE during daily activities. The CIMT group reported using the more affected UE for an average of 14 activities and the traditional intervention group for an average of 15 activities before treatment and both groups for 24 after treatment.

**DISCUSSION**

This study is, to our knowledge, the first RCT to use kinematic analysis to investigate differences in motor-control strategies after CIMT versus traditional intervention. We observed better performance in reaching kinematics of the UE after CIMT as compared with the traditional intervention. The CIMT group also showed less motor impairment assessed by the FMA and greater functional ability evaluated by the MAL. The beneficial effects of CIMT are consistent with previous findings,7-9,11-13 but the findings further suggest that CIMT may improve motor-control strategies evidenced by kinematic data.

The kinematic results of this study extend those of the case report by Hakim et al23 to a larger sample. The enhanced

Table 1: Subject Clinical and Demographic Characteristics

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>CIMT (n=24)</th>
<th>TI (n=23)</th>
<th>( P^* )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex (male/female)</td>
<td>16/8</td>
<td>17/7</td>
<td>.83</td>
</tr>
<tr>
<td>Mean age ( \pm ) SD (y)</td>
<td>53.93( \pm )11.20</td>
<td>56.77( \pm )12.90</td>
<td>.42</td>
</tr>
<tr>
<td>Side of lesion (right/left)</td>
<td>11/13</td>
<td>11/12</td>
<td>.89</td>
</tr>
<tr>
<td>Mean months after stroke ( \pm ) SD</td>
<td>12.51( \pm )9.64</td>
<td>11.98( \pm )11.72</td>
<td>.87</td>
</tr>
<tr>
<td>Between 0.6 and 6m (n)</td>
<td>9</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Between 6 and 12m (n)</td>
<td>5</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Between 12 and 24m (n)</td>
<td>6</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Between 24 and 37m (n)</td>
<td>4</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Brunnstrom stage (proximal part of UE) (median)</td>
<td>4.5</td>
<td>4.5</td>
<td>.39</td>
</tr>
<tr>
<td>Mean modified MMSE score ( \pm ) SD</td>
<td>84.42( \pm )10.35</td>
<td>84.43( \pm )11.62</td>
<td>.99</td>
</tr>
<tr>
<td>Mean pretreatment performance on kinematic variables ( \pm ) SD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reaction time (s)</td>
<td>0.72( \pm )0.46</td>
<td>0.60( \pm )0.28</td>
<td>.28</td>
</tr>
<tr>
<td>Normalized movement time</td>
<td>0.07( \pm )0.05</td>
<td>0.09( \pm )0.06</td>
<td>.95</td>
</tr>
<tr>
<td>Normalized total displacement</td>
<td>1.70( \pm )0.51</td>
<td>1.42( \pm )0.36</td>
<td>.036</td>
</tr>
<tr>
<td>Peak velocity (cm/s)</td>
<td>66.32( \pm )22.88</td>
<td>55.35( \pm )15.45</td>
<td>.062</td>
</tr>
<tr>
<td>Normalized movement units</td>
<td>7.06( \pm )5.26</td>
<td>5.69( \pm )6.25</td>
<td>.42</td>
</tr>
<tr>
<td>Mean pretreatment scores on FMA ( \pm ) SD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UE</td>
<td>39.50( \pm )13.45</td>
<td>41.74( \pm )13.47</td>
<td>.64</td>
</tr>
<tr>
<td>AOU</td>
<td>0.64( \pm )0.86</td>
<td>0.60( \pm )0.92</td>
<td>.87</td>
</tr>
<tr>
<td>QOM</td>
<td>0.72( \pm )1.01</td>
<td>0.69( \pm )1.17</td>
<td>.94</td>
</tr>
</tbody>
</table>

Abbreviations: MMSE, Mini-Mental State Examination; SD, standard deviation; TI, traditional intervention.

\( * \)P associated with the chi-square test for categorical variables, the independent t test for continuous variables, and the Mann-Whitney U test for the ordinal variable.

---

Table 2: Descriptive and Inferential Statistics for Analysis of Reaching Kinematics and Clinical Assessment

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Posttreatment</th>
<th>ANCOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CIMT (n=24)</td>
<td>TI (n=23)</td>
</tr>
<tr>
<td>Kinematic variables</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reaction time (s)</td>
<td>0.48( \pm )0.17</td>
<td>0.63( \pm )0.32</td>
</tr>
<tr>
<td>Normalized movement time</td>
<td>0.04( \pm )0.03</td>
<td>0.05( \pm )0.03</td>
</tr>
<tr>
<td>Normalized total displacement</td>
<td>1.32( \pm )0.22</td>
<td>1.42( \pm )0.39</td>
</tr>
<tr>
<td>Peak velocity (cm/s)</td>
<td>76.42( \pm )16.17</td>
<td>65.72( \pm )18.79</td>
</tr>
<tr>
<td>Normalized movement units</td>
<td>0.13( \pm )0.11</td>
<td>0.19( \pm )0.16</td>
</tr>
<tr>
<td>FMA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UE</td>
<td>46.75( \pm )11.58</td>
<td>44.78( \pm )13.08</td>
</tr>
<tr>
<td>MAL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AOU</td>
<td>1.85( \pm )1.24</td>
<td>0.81( \pm )1.13</td>
</tr>
<tr>
<td>QOM</td>
<td>1.85( \pm )1.14</td>
<td>0.84( \pm )1.08</td>
</tr>
</tbody>
</table>

Arch Phys Med Rehabil Vol 88, August 2007
performance in reaching kinematics (ie, movement quality) of the CIMT group may reflect better control strategies (ie, a more preprogrammed strategy) as a result of forced use of the more affected UE for intensive practice on functional tasks. The tasks offered motor problems to be solved and allowed subjects to experiment on the solutions to accomplish the task goal. Moreover, the decreased standard deviations for kinematic variables imply a more stable movement pattern in the CIMT group. Such better performance indicates that subjects who had received CIMT were able to perform more preprogrammed movement than those who had received the traditional intervention. In other words, subjects who had received CIMT used their more affected UEs more spontaneously and could modify ongoing movement with less attentional effort. Additional evidence for better movement control after CIMT was evidenced by faster movement planning or preparation (less reaction time), increased movement efficiency (less normalized movement time), and a straighter and smoother movement trajectory (less normalized total displacement and fewer movement units). This enhanced movement control may have related to short-term learning changes at central and spinal levels.36 Intensive training of the more affected UE may, at the central level, enhance motor planning and at the spinal level decrease the latency between activation of agonist and antagonist muscles, leading to a shorter reaction time.37 Extensive practice with a variety of functional tasks may have provided opportunity for the patients to experiment with efficient arm use for reaching movements and regain motor skills and thus perform the reaching task faster and more efficiently. Straighter and smoother movements might be caused, at the central level, by improved motor planning involving interjoint coordination.38 Improved coordination of movements might also, at the spinal level, be caused by an increase in the intensity of activation of spinal motoneuron pools, leading to increased efficiency and coordination of muscular contraction.36

The difference in peak velocity posttreatment was not significant between CIMT and traditional intervention. One possibility pertains to the large variability in peak velocity within both groups. Another possibility is that peak velocity is an integral factor involving both spatial and temporal properties of movements and is associated with force or impulse at movement initiation.32 Because neither treatment program in our study emphasized muscle strengthening or force control, force-related performance during reaching movements may be considered a less sensitive variable for detecting differences in therapeutic effect. When improvement of movement velocity is an endpoint, treatment should incorporate tasks demanding force control (eg, turning a key to open a door). Future research might also evaluate force-related performance during speed-emphasized reaching tasks.

The enhanced performance in movement kinematics associated with CIMT suggests that the therapeutic gains may reflect reduced deficits in motor-control strategies. Movement information on control strategies derived from kinematic analyses may complement outcomes of impairment and functional performance derived from clinical measures and may allow a better understanding of motor-control mechanisms underlying CIMT. This knowledge about how patients with stroke perform reaching movements provides unique perspectives on characteristics of control strategies and may contribute to outcome evaluations in stroke rehabilitation. It is also important for future research to use kinematic analysis to examine motor-control strategies in a variety of functional UE tasks to investigate the generalizability of the current findings to different types of daily activities. Several neuroimaging studies39-43 have shown that CIMT leads to cortical reorganization with extension, shift, and recruitment of ipsilesional, perilesional, and contralesional cortical areas of the sensorimotor network. It remains unclear whether the therapeutic benefit is attributable to neuroplastic changes in brain areas subserving activity-dependent cortical reorganization. A possibility for future research is to use functional magnetic resonance imaging analysis pre- and post-CIMT to investigate the possible relation between kinematic improvement of reaching movements and altered brain activation.

The finding of better performance represented by FMA and MAL scores after CIMT versus traditional intervention may reflect less motor impairment and increased use of the more affected UE after CIMT. The finding on the FMA is consistent with most previous studies7-8,12,14 but not that of van der Lee et al.3 This discrepancy might be primarily caused by differences in the treatment program for the comparison group. The present study, along with those of Page,7,8,12 and Boake14 and colleagues, used the traditional intervention for the comparison group, whereas the study by van der Lee3 used bimanual task training for the comparison group. Bimanual task training involved mass practice of both hands, resulting in only slightly greater reduction of motor impairment in the CIMT versus comparison group. The findings on the MAL are consistent with previous findings7,9,12,14 and suggest that the learned nonuse phenomenon observed in the patients with stroke can be overcome through CIMT.

This study recruited patients with various levels of arm-movement impairment ranging from hemiplegia with persistent synergy patterns but preservation of some isolated arm movements to hemiparetic weakness with isolated finger movements. The application of CIMT might not be limited to the high-functioning patients having finger movements and might be extended to the patients having only some isolated arm movements. The stage of recovery from stroke in subjects of this study spanned from acute, to subacute, and to chronic. The issue that the natural recovery of stroke might confound the observed effects of CIMT was taken into account in this study. There were almost equal numbers of subjects at each recovery stage between the 2 groups. The variable of time poststroke was further addressed in the statistical analysis by treating it as a covariate. The confounding effect of natural recovery is thus not a plausible explanation for the beneficial effects of CIMT.

Study Limitations

There is a noteworthy limitation of our study. The instruments used in the study assessed control strategies inside the laboratory. Future research might determine if the therapeutic effect of CIMT transfers to the nonlaboratory environment. The absence of electromyographic data in the current study is another limitation; evaluation of muscle-activation patterns together with evaluation of spatiotemporal-movement control might provide a more complete picture of the specific movement parameters that may be affected by CIMT. A clearer understanding of which movement parameters are responsive to therapeutic change and the nature of such change can provide insights into the mechanisms responsible for the therapeutic effects on control strategies. As a further limitation, the long-term effects of CIMT on kinematic performance are unclear and await longitudinal study with long-term follow-up.44 Future research may also study whether improvements in control strategies after CIMT predict better motor relearning abilities, more recovery in daily functions, and increased potential for return to work.
CONCLUSIONS

This RCT used kinematic and clinical analyses to study postintervention differences between CIMT and traditional intervention in stroke patients varying in the level of arm motor impairment. The study showed CIMT improved both kinematic performance (except for peak velocity) at the level of control strategy and motor performance at the impairment and functional levels. Further research may use brain-imaging techniques to elucidate whether improved control strategies after CIMT are associated with treatment-induced cortical reorganization.

References


Suppliers

a. Model 370; Vicon, 14 Minns Business Pk, West Way, Oxford OX2 0B, UK.

b. National Instruments Corp, 11500 N Mopac Expwy, Austin, TX 78759-3504.
A Randomized Controlled Trial of an Implantable 2-Channel Peroneal Nerve Stimulator on Walking Speed and Activity in Poststroke Hemiplegia

Anke I. Kottink, MSc, Hermie J. Hermens, PhD, Anand V. Nene, MD, PhD, Martin J. Tenniglo, PT, Hans E. van der Aa, MD, PhD, Hendrik P. Buschman, PhD, Maarten J. IJzerman, PT, PhD


Objective: To determine the effect of a new implantable 2-channel peroneal nerve stimulator on walking speed and daily activities, in comparison with the usual treatment in chronic stroke survivors with a drop foot.

Design: Randomized controlled trial.

Setting: All subjects were measured 5 times in the gait laboratory.

Participants: Twenty-nine stroke survivors with chronic hemiplegia with drop foot who fulfill the predefined inclusion and exclusion criteria were included in the study.

Intervention: The intervention group received an implantable 2-channel peroneal nerve stimulator for correction of their drop foot. The control group continued using their conventional walking device, consisting of an ankle-foot orthosis, orthopedic shoes, or no device.

Main Outcome Measures: Walking speed, assessed both by a six-minute walk test (6MWT) and by using a 10-m walkway, was selected as primary outcome measure and activity monitoring data, consisting of percentage time spent on stepping, standing, and sitting/lying were selected as secondary outcome measure.

Results: Functional electric stimulation (FES) resulted in a 23% improvement of walking speed measured with the 6MWT, whereas the improvement in the control group was only 3% (P=.010). Comfortable walking speed measured on a 10-m walkway was also significantly improved in favor of FES (P=.038). The percentage time spent on stepping deteriorated with 3% in the intervention and 0.8% in control group, which was not statistically significant between both groups (P=.13).

Conclusions: The present study shows a clinically relevant effect of the implantable 2-channel peroneal nerve stimulator on walking speed in the sample of stroke survivors included in our study.

Key Words: Electric stimulation; Foot; Peroneal nerve; Rehabilitation; Stroke; Walking

Foot Drop or Drop Foot is a simple term that describes a rather complex problem. A variety of conditions, such as dorsiflexor injuries, peripheral nerve injuries, stroke, neuropathies, drug toxicities, or diabetes can be associated with drop foot. Depending on the cause, drop foot may be temporary or permanent. It can be defined as a significant weakness or absence of ankle and toe dorsiflexors. These muscles assist in clearing the foot during swing phase and control plantarflexion of the foot on heel strike. Weakness or absence of this group of muscles associated with imbalance between invertors and evertors results in an equinovarus deformity. Walking becomes a challenge due to the patient’s inability to control the foot at the ankle. In fact, due to predominance of extensor synergy, hip and knee flexion are usually both reduced which further lengthens the limb functionally. Accordingly, many stroke survivors use circumduction and hip hiking and on occasion vaulting in order to compensate.

This article focuses on drop foot in subjects with chronic hemiplegia after stroke. Drop foot after stroke is thought to be caused partly by poor active control of the anterior tibial muscle and by increased and inappropriate tone in the muscles of the leg, particularly the calf. Functional electric stimulation (FES) is the clinical application of electric current to the intact nerves of the body, in order to generate a muscle contraction. This contraction is then incorporated into a functional activity, for example, walking. FES systems for the treatment of drop foot are in clinical use in significant numbers, especially the surface Odstock Drop Foot Stimulator. These studies showed an increased walking speed and a reduction in Physiological Cost Index (PCI), which is a measure for energy cost. The perception of the users was that the Odstock stimulator was of clinical benefit and therefore the compliance was very high.

In a systematic review, the results of 8 studies were analyzed to assess the orthotic effect of FES on walking in stroke survivors with a drop foot. The pooled effect size for walking speed was .13m/s (range,.07-.20m/s) or 38% (range, 22.18%-53.8%). FES also seemed to have a positive orthotic effect on the PCI. Recently, another meta-analysis published where the therapeutic effect of FES on walking speed in stroke survivors was determined. A significant mean difference in walking speed of .18m/s was found, indicating the effectiveness of FES treatment.

At present, only 1 randomized controlled trial (RCT) examined the effect of common peroneal nerve stimulation on walk-
ing speed in stroke survivors with chronic hemiplegia. The study reported a significant improvement of 20.5% in walking speed in the FES group, whereas the control group, who received physiotherapy (PT) only, showed a nonsignificant improvement of 5.2%. However, this RCT did not use an ankle-foot orthosis (AFO) as the control device. A comparison was made between walking with FES versus no treatment. In clinical practice, the conventional treatment of a drop foot is an AFO, which is most often a plastic support worn in the shoe to keep the ankle in a neutral position. An interesting aspect that has not been clarified yet is the additional value of the peroneal nerve stimulator in comparison with an AFO as the control device.

Surface-based FES is the common approach in the clinical setting, but there are several problems with this approach including difficulty with electrode positioning and skin allergy. Assuming that the drop foot requires a permanent solution, an implantable system might be considered. Potential advantages include stability of electrode position, easier donning and doffing of the system, and reduced pain and skin irritation. Several implantable systems have been developed in the past. These were 1-channel stimulators, which did not allow for differential activation of peroneus and anterior tibial muscles for inversion-eversion balance post surgery. Accordingly, an implantable device with 2 independent channels was developed.

The aim of the present RCT was to determine the effect of an implantable 2-channel peroneal nerve stimulator on walking speed and physical activity in comparison with the usual treatment in stroke survivors with a drop foot. We hypothesized that the intervention group would improve their walking speed by at least .20m/s, defined by Perry et al as clinically relevant over a device use period of over 6 months. Furthermore, we expected that the intervention group would show an increase in physical activity at the end of the trial. No changes in the control group were expected in both outcome measures.

METHODS

Study Design

We conducted this study as an RCT and the CONSORT statement was used to report the trial. All subjects were assessed 5 times in the gait laboratory. The baseline measurement took place about 1 week before the randomization procedure. The follow-up measurements were performed 4, 8, 12, and 26 weeks after the surgical procedure in the intervention group. Subjects assigned to the control group were measured in the same weeks as subjects assigned to the intervention group.

Walking speed, measured both by the six-minute walk test (6MWT) and on a 10-m walkway, was defined as the primary outcome measure. The 6MWT was measured only during baseline, week 12, and week 26. Comfortable walking speed, measured on a 10-m walkway, was measured during all assessments. Physical activity, the secondary outcome measure, was measured at baseline and at week 26 by monitoring a randomly selected group of subjects from both the intervention and control groups. Unfortunately, it was not possible to monitor all participating subjects, because only 2 acti6 PAL systems were available during the trial.

All measurements were performed by the same examiners (AVM). Instructing the subjects in the intervention group on the proper use of the peroneal nerve stimulator and assessment of stimulation levels of the 2 output channels took place on the same day as the outcomes assessment. This was done for 2 reasons: (1) to keep the number of visits similar for both study groups so that the same amount of attention was paid to both groups; and (2) to save time and travel costs in the intervention group. If problems were experienced by the patients they were instructed to report them immediately, so that they could be resolved as soon as possible. Blinding of both the study personnel and participants was not possible due to the surgical procedures. All data were analyzed according to an intention-to-treat principle, which means that all participants in the trial were analyzed according to the treatment to which they were allocated, whether they received it or not.

Participants

Most subjects were recruited to the trial in response to an article in a local Dutch newspaper about the first results with the implantable stimulator in stroke survivors. Some patients were recruited through consultant and general practitioner referrals and in some cases on the advice of the physiotherapist treating the patient. Both recruitment strategies resulted in a selected sample of well-motivated stroke survivors. Because of the invasive nature of the treatment, selection of the most appropriate patients is very important. Therefore patients that were most likely to experience treatment success were selected. Table 1 shows the criteria for inclusion and exclusion in the trial.

<table>
<thead>
<tr>
<th>Inclusion Criteria</th>
<th>Exclusion Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drop foot identified by an inability to achieve a normal heel strike during walking</td>
<td>Age &lt; 18y</td>
</tr>
<tr>
<td>First hemiplegia of at least 6 months in duration as a result of a CVA with a stable neurology</td>
<td>Passive dorsiflexion of the ankle &lt; 5° with knee in extension</td>
</tr>
<tr>
<td>Subject is an outdoor walker</td>
<td>Medical conditions other than CVA, that is, neurologic, rheumatic, cardiovascular, or systemic disorders (including diabetes mellitus) limiting the function of walking</td>
</tr>
<tr>
<td>Able to give an informed consent</td>
<td>Injury to deep and superficial peroneal nerve and sciatic nerve</td>
</tr>
<tr>
<td></td>
<td>Any medical condition that would exclude the use of a surgical procedure or anesthetic</td>
</tr>
<tr>
<td></td>
<td>Not able to don and doff the equipment</td>
</tr>
<tr>
<td></td>
<td>Pregnancy</td>
</tr>
</tbody>
</table>

Abbreviation: CVA, cerebrovascular accident.
Subjects that fulfilled the selection criteria were admitted to the trial. One week after the baseline assessment, subjects were randomly allocated to either the intervention group or control group. Random allocation was performed in blocks of 2 subjects, to ensure a close balance of the numbers in each group. The randomization procedure was performed by an independent person. All patients were allowed to continue their usual PT sessions during the trial, which was recorded in a patient diary.

The experiment protocol was approved by the local medical ethics committee and the subjects gave their informed consent before participating in the experiments.

The Stimulation System

The implantable 2-channel peroneal nerve stimulator consists of an external transmitter with a built-in antenna, a foot switch, and implantable components consisting of the stimulator, the 2 leads, and the bipolar intraneural electrodes (fig 1).10,15

Transmitter. The transmitter uses a single 40-mm diameter transmission coil that transmits alternately on 2 frequencies. This switching results in a pulse repetition rate of 30Hz on each channel. The amplitudes of the monophasic pulses modulated on each carrier wave are controlled separately. The transmitter weighs approximately 0.1kg and is attached with a strap on the lateral side of the lower leg, over the site of the implant, just below the knee. A footswitch placed under the heel of the patient’s foot inside the shoe determines the on-and-off switching of the stimulation. The transmitter battery is charged overnight.

Implanted stimulator. The implantable 2-channel peroneal nerve stimulator is a passive device, receiving information carried by the radiofrequency signals and converting them into the stimulation pulses of the desired amplitude and frequency. The receiver block is approximately 33mm in diameter and 6mm thick. It contains 2 independent and galvanically separate electric circuits built on a ceramic substrate 29mm in diameter. The 2 circuits are tuned to operate at different frequencies, namely, 1 and 2MHz, allowing them to be individually controlled by the transmitter. This further reduces the risk of cross-talk between the channels. The electronic circuits of the receiver block are encapsulated in silicone rubber elastomer.

The electrodes are placed under the epineurium of the nerve, providing good mechanical stability and, due to the proximity to the nerve fascicles, low stimulation currents can be used. The 2 cables that connect the receiver block to the electrode arrays are composed of 2 helically wound platinum wire conductors with enamel insulation. The intertwined helixes are encased in silicone rubber elastomer. The electrodes are surgically positioned at 2 distinct locations, which are determined during test stimulation with a hook electrode combined with visual inspection of the generated movement. One electrode is placed under the epineurium of the superficial peroneal nerve (SPN) (eversion) and the other under the epineurium of the deep peroneal nerve (DPN) (dorsiflexion). The stimulation pulses have an asymmetric biphasic charge balanced waveform.

Surgical procedure. For patients receiving the implant, we performed a presurgical nerve conduction velocity measurement to check the integrity of the deep and superficial branches of the peroneal nerve. The surgical procedure was performed under general or spinal anesthesia. The patient’s knee was placed in a flexed position to relax the common peroneal nerve (CPN). An incision of approximately 50mm along the course of the CPN was made, beginning just below the head of the fibulae. The CPN and its 2 branches, the DPN and the SPN, were visually identified. Nerve identity was checked using a surgical nerve stimulator and hooked electrodes. The response to stimulation of each nerve was checked at several stimulation sites, along and around the nerve. When the site for optimal responses was identified on both nerves during test stimulation, a small incision through the epineurium at each of the 2 nerves was made. At each of the 2 sites, an electrode was inserted.
underneath the epineurium, which was fixed in place using sutures or tissue col fixative, a surgical glue. Using tags, positioned approximately 15mm from the end of the electrode, the leads were sutured to the fascia of the underlying tibialis muscle. The leads were arranged in such a way as to reduce as far as possible the possibility of mechanical loading on the electrode sites resulting from muscle activity. The receiver body was placed in a subcutaneous pocket. Several test stimulations took place before closing the wound.

Two weeks after the surgery the wound was checked and a first test stimulation took place. In the third week, stimulation during walking was tested and the stimulator was taken home by the patient. The use of the stimulator was gradually increased over 2 weeks to prevent severe muscle pain and fatigue. After this period patients were allowed to use the system all day.

Measurements

Assessment of 6MWT. The 6MWT was used to estimate the walking speed during daily activity. Butland et al. reported that results of the 6MWT are highly reproducible and show moderate to strong correlations with comparable outcome measures. An oval course with known distance was clearly defined in the gait laboratory. The distance walked in 6 minutes at a comfortable walking speed was recorded using a stopwatch. Subjects were allowed to use a walking stick if necessary and this was recorded. The condition during baseline was the standard for the follow-up measurements. During baseline, all subjects walked with their conventional walking aid for the correction of their drop foot. During the follow-up measurements, the control group walked with their conventional walking aid and the intervention group walked with the implantable 2-channel drop foot stimulator. Walking speed was calculated by dividing the walking distance by 360 (6min = 360s).

Assessment of walking speed. We instructed patients to walk at a comfortable walking speed in the gait laboratory on a 10-m walkway; no other instructions were given. Walking speed was measured automatically by using the Vicon system, consisting of 2 infrared beams over a distance of 7.5m. To exclude the influence of acceleration and deceleration at the start and finish of the walkway, 1.5m were allowed at the beginning and end of the walkway. Walking speed measured on a 10-m walkway was found to be a valid, reliable and responsive outcome measure. Control subjects were asked to perform the walk 4 times without and 4 times with their conventional walking aid during all measurements. At baseline, intervention group subjects were asked to walk 4 times without and 4 times with their conventional walking aid. During the follow-up measurements, they were asked to walk 4 times without and 4 times with stimulation. All subjects were allowed to use a walking stick if needed and this was recorded. The condition during baseline was the standard for the follow-up measurements. In each walking condition, the first walk was excluded from analysis. For each walking condition a mean walking speed was calculated by averaging the 3 remaining walking sessions.

Assessment of activity level. We used the activPAL professional to electronically monitor the level of activity in patients’ home environment. The activPAL is an accelerometer based measurement device used to record subjects’ primary physical activities (stepping, standing, sitting/lying) during their daily life. Data from the activPAL have been shown to be both valid and reliable. The device was fixed using an adhesive tape on the mid-line of the thigh, midway between hip and knee.

Because only 2 activPAL systems were available during the trial, it was not possible to monitor all participating subjects. Therefore, 10 subjects of the control group and 11 subjects of the intervention group were randomly selected to be monitored over a 5-day period, with exception of the weekend, through all waking hours during 2 evaluation periods. The first evaluation period was performed at baseline and the second evaluation period was performed at week 26. The selected outcome parameters were the percentage time spent on stepping, standing, and sitting/lying. Mean values were calculated by averaging the values found on the 5 recorded days.

Statistical Analysis

We performed a power analysis based on estimates that were obtained from the pooled analysis in a previous review from our group. The review intended to analyze the orthotic effect of FES on the improvement of walking in stroke patients with a drop foot. Data on walking speed of 4 clinical papers were pooled to estimate a mean difference ± standard deviation (SD) of 0.134±0.124m/s. The following numbers were used for the power analysis: mean 1, .000; 1.124; mean 2, 11.34±124; δ=.134; α=.05; and power, .80. The power calculation resulted in a number of 14 subjects in each group.

Baseline characteristics of the 2 groups were compared to evaluate the success of randomization. Walking speed parameters were tested for normality using the Shapiro-Wilks test, indicating a normal distribution for walking speed values (P=.05). We used linear mixed-model analyses to determine the overall orthotic effect of FES on both walking speed parameters when compared with the conventional treatment. An advantage of this method is that all available data could be included in the analysis, even if some data were missing. Group (FES, conventional treatment), time outcome assessments (−6, 4, 8, 12, 26wk), and the interaction between group and time were entered as terms in the model. The interaction is used to test differences between both groups in the change in outcome measured over time. Differences between and within both groups over the period between baseline and week 26 were evaluated. The model was also used by us to measure the strength of association between both walking parameters. Post hoc tests were performed for both walking speed parameters with Sidak-adjusted multiple comparisons.

Because of the non-normal distribution of the activ PAL data (Shapiro-Wilks test, P=.05), we used the Wilcoxon signed-rank test to compare the activ PAL data in both groups between both evaluation periods. The significance level α was set at .05 for all tests. All statistical analyses were performed with SPSS for Windows.

RESULTS

Participants

Figure 2 shows the flow of participants throughout the study. Twenty-nine patients with drop foot due to cerebrovascular accident participated in the present study. Table 2 shows the characteristics of the participants. Subject characteristics were not very different between both groups, with the exception of mean time after stroke, which was longer for the intervention group. At baseline both groups also exhibited no significant difference in walking speed, assessed both by a 6MWT and by using a 10-m walkway, indicating that the intervention and control group were matched.

In the intervention group 7 patients used a plastic AFO, 3 patients wore orthopedic shoes and 4 patients did not use a walking aid to correct their drop foot just prior to the start of
the trial. In the control group all subjects wore a plastic AFO. Four subjects dropped out of the study, 1 woman in the intervention and 3 men in the control group. The implant of the subject in the intervention group who dropped out failed after functioning properly for about 10 weeks. An investigation of the explanted system showed that the failure was caused by a technical defect in the epineural electrode responsible for the dorsiflexion movement. Two subjects in the control group withdrew after the randomization procedure and the other subject dropped out in week 11, because of psychologic issues not related to the study. The remaining subjects in the intervention group did not report any technical failure of the stimulation system and continued to use the stimulator during the entire follow-up period.

At baseline, walking speed when using the conventional walking aid in the intervention group was determined by means of calculating the mean walking speed for the subjects who used a plastic AFO or orthopedic shoes (n=10).

### Six-Minute Walk Test

Figure 3 shows both the mean walking distance that was reached during the 6MWT in both groups and the calculated walking speeds. Because no difference in baseline values was present between both groups, no correction for baseline value was necessary in the analysis.

A significant difference between both groups was found when all assessments were taken into the linear mixed model (P=.010), showing a positive effect of FES on the performance of the 6MWT. At the first follow-up assessment (week 12), both the intervention and control group showed an improvement in the performance of the 6MWT. However, post hoc analysis showed that the change in walking speed at 12 weeks relative to baseline did not differ significantly between groups (P=.49). At 26 weeks, the intervention group continued to show improvements, whereas the control group exhibited some deterioration. Post hoc analysis showed that the change in
walking speed at 26 weeks relative to baseline now differed significantly between groups \((P = .049)\).

**Walking Speed**

Figure 4 shows the results of the assessments for both groups obtained on all different walking speed conditions. Because there was no difference in baseline values, correction for baseline value was not included in the analysis.

When no walking device was used, no significant difference in walking speed between groups was found when all assessments were taken into the linear mixed model \((P = .152)\). The changes within both the intervention and the control group over time, relative to their baseline values, were also not statistically significant \((P = .812, P = .112\), respectively).

When the control group used their walking aid and the intervention group used their FES, the linear mixed model indicated significant differences between groups \((P = .038)\). Walking speed remained constant over time within the control group \((P = .572)\). The intervention group showed a small deterioration in walking speed immediately after starting with the FES treatment (week 4), followed by an improvement in walking speed when FES was used for a longer period. Overall, when baseline was compared with the last follow-up assessment, the change in walking speed within the intervention group over time was statistically significant \((P = .01)\).

When comparing only the last follow-up assessment with baseline between both groups, a trend toward statistically significant effect of FES on walking speed over time was found in comparison with the conventional treatment \((P = .097)\). At none of the follow-up assessments, where the use of FES in the intervention group was compared with the use of an AFO in the control group, did post hoc testing result in a statistically significant difference relative to baseline.

To examine the relation between both measurement techniques used to obtain comfortable walking speed, correlation coefficients were calculated. Both the 6MWT and the 10-m walkway were performed during baseline, week 12, and week 26. The correlation coefficients were around .90 for all 3 assessments, indicating a strong relation.

**The activPAL Professional**

Table 3 shows the activPAL data measured in both groups. From the table it can be seen that the percentage time spent stepping deteriorated 3% and 0.8% in the intervention and control groups, respectively, when the first evaluation period was compared with the second. However, the difference between groups was not statistically significant \((P = .13)\). The percentage time spent standing declined approximately 3% in the intervention group, whereas the control group improved 2%. There was a trend toward statistically significant difference between groups \((P = .06)\). Time spent on sitting/lying increased approximately 6% in the intervention group, whereas the control group declined approximately 1%. The difference was statistically significant \((P = .04)\).

**DISCUSSION**

The primary aim of this study was to determine the effect of using an implantable 2-channel peroneal nerve stimulator on comfortable walking speed in comparison with the usual treatment in stroke survivors with a drop foot. Walking speed was measured in 2 different ways: by measuring average speed during a 6MWT and by measuring speed on a 10-m walkway.

The results of the present study show that FES results in a significant improvement in walking speed measured with the 6MWT when FES is used for a period of about 6 months as a treatment for drop foot in chronic stroke survivors. Furthermore, comfortable walking speed measured on a 10-m walkway also increases significantly in the intervention group with regard to the control group during the trial when all follow-up assessments are taken into the analysis.

Walking speed was measured in 2 different ways. Some differences might have been expected. The patient may feel greater pressure to perform well during the 10-m walkway, whereas the longer duration of the 6MWT might reflect the more natural cadence and velocity. Another possible source of difference might be the longer duration of the 6MWT, which might induce a fatigue effect. However, both measurements gave very similar results; no systematic differences were found.

In the systematic review of Kottink et al., in which the effect of FES on walking speed in stroke survivors with a drop foot was evaluated, most of the included studies did not compare FES with the conventional orthotic treatment, but made a comparison between the conditions with and without FES. We intended to perform a more pragmatic trial. The present study is only the second RCT done on drop foot stimulation and the first RCT that examined the effect of an implantable drop foot stimulator.

All subjects included were stroke survivors that were in a chronic phase, so spontaneous recovery was not expected to be a confounder in the present study. This is confirmed by obvious lack of changes in the walking speed in the control group over time. The improvements measured in the intervention group can therefore be completely attributed to the FES treatment.

Another characteristic of our patient group was that they all had a relatively good walking function at the start of the trial, which is a result of our strict predefined inclusion and exclusion criteria. Richards et al. described that to be independent
in the community, a speed of .80±.18m/s is required. When looking at the walking speeds measured at baseline, one can conclude that all subjects in the trial satisfied this criterion. Most studies that examined the effect of peroneal nerve stimulation included subjects with a more impaired walking function. The control and intervention groups in the study by Burridge et al7 walked with a speed of .48 and .64m/s, respectively, at baseline. Patients included in the study by Bogataj et al,20 consisting of acute, subacute, and chronic stroke survivors, showed a mean walking speed of .19 and .23m/s for FES and control groups, respectively. In the literature, Wielers et al43 described that FES systems were of most benefit to subjects who walked very slowly. They explained that the smaller improvements among less impaired walkers were due to the fact they already had good control over many muscle groups. Consistent with this hypothesis, Ladouceur and Barbeau22 reported a negative correlation between initial walking speed and the effect of FES. Thus, the inclusion of more severely impaired patients may be associated with a larger treatment effect. Although implantable stimulators have clear advantages with respect to the accuracy of stimulation and user comfort, it is also obvious that surgery is required and overall costs are considerably higher. Thus, candidates for the implantable system should first be given a trial period with a surface peroneal nerve stimulator to assess and appreciate the potential benefits of using electric stimulation. When these users then encounter specific problems that might be amenable to implantable systems such as poor electrode reliability, painful sensation, and difficulty with donning and doffing the system, implantation should be considered. A slight deterioration in walking speed measured on a 10-m walkway was seen in the intervention group at the first follow-up. This can be explained by the anticipated inactive period and deconditioning after the implantation procedure. As the deterioration was observed with and without FES use, this explanation appears plausible. At the last follow-up assessment the intervention group continued to show improvements in walking speed, which suggests that a plateau has not been reached.

Study Limitations

A limitation of this study is the small sample size. This is reflected in the finding that FES resulted in a significant improvement of comfortable walking speed measured on a 10-m walkway when all assessments were included in the linear mixed model (P=.038), but only a trend toward significance when baseline was compared with the last follow-up assessment (P=.097). The more data are included in the linear mixed model, the higher the power. Reaching sufficient power is often reported as a problem in studies performed in a rehabilitation setting.

Clinical Relevance

An interesting aspect to discuss is the clinical relevance of the results found in the present study. Clinicians can use this information to determine the effectiveness of the FES treatment. Perry et al11 reported that a difference of .20m/s in walking speed with and without AFO was defined as clinically relevant. From figure 4 it can be seen that at the final follow-up assessment the control group shows a difference of .07m/s in walking speed measured on a 10-m walkway between the walking conditions with and without conventional walking aid. A difference of .21m/s is found in the intervention group when walking with FES was compared with walking without FES, which is clinically relevant in accordance with Perry. Because the 6MWT was only performed while using a walking aid in both groups, it was not possible to test if these walking speed results were clinically relevant in accordance to Perry. However, another definition of clinical relevance was given by Burridge,7 who considered a percentage change of 10% in walking speed to be functionally relevant. When looking at the walking speeds converted from the 6MWT, the intervention and control groups show an improvement of 23% and 3%, respectively. Thus, according to Burridge,7 the results found in the intervention group are highly clinically relevant. When the comfortable walking speed results measured on a 10-m walkway are taken into consideration, an improvement of exactly 10% is found in the intervention group when using the FES system (baseline, .80m/s; week 26, .88m/s). The control group did not show a change in walking speed during the trial when using their walking device. During both baseline and the last follow-up assessment walking speed was .74m/s. Some studies in the literature reported the same amount of improvement in walking speed measured on a 10-m walkway as was found in the present study. Waters et al,8 who also studied the effect of an implantable stimulator in chronic stroke survivors, found a difference of .24m/s in free cadence walking speed when walking with FES after surgery compared to walking without an orthosis before surgery. However, no control group was included in their study. The intervention group in the study performed by Bogataj20 showed an increase of .22m/s in walking speed when FES was compared with walking without FES, whereas the control group improved .03m/s. Our hypothesis that the intervention group would show an improvement in walking speed that is clinically relevant is confirmed by the present study results.

The secondary aim of the present study was to measure physical activity by using the activPAL professional. Physical activity significantly decreases in the intervention group, as indicated by the decrease of time spent on stepping and more time spent on sitting/lying. This finding is remarkable as one would expect an increase of time spent on stepping, induced by the greater walking speed enabling people to walk greater distances.

Recently, Stein et al22 found that with surface stimulation the number of steps per day, as reflected in the number of delivered pulse trains, increased significantly over time. However, a direct comparison with the present study is hampered, because we investigated physical activity of subjects while using their
conventional walking aid at baseline in comparison with physical activity while using the implantable device at week 26. Because all subjects in our study were in the chronic phase of stroke, they were all trained well in walking with their conventional walking aid. Physical activity at baseline is therefore expected to be at a higher level in the present study, making it more difficult to find differences in due time. Our results do suggest that patients do not change their averaged walking distance using FES whereas the increase in walking speed allows them to spend even less time in walking. Despite the decrease measured in daily activity, the overall results suggest that an implanted peroneal nerve stimulator is an effective treatment option for a select group of chronic stroke survivors with foot drop.

CONCLUSIONS

FES resulted in a significant and clinically relevant increase in walking speed. The intervention group showed an improvement of 23% in walking speed in comparison with a 3% improvement in the control group when walking speed was measured by means of the 6MWT. Comfortable walking speed, measured on a 10-m walkway, was also significantly improved in favor of FES. In contrast to our expectations, the results found by the activPAL professional do suggest that the average walking distance did not change by applying the stimulator.

In conclusion the results suggest that the implantable 2-channel peroneal nerve stimulator is a clinically relevant treatment option in a select group of chronic stroke survivors. Future studies might investigate the generalizability of the results to other stroke survivors and the relevance of the stimulation system in other patient categories with upper motoneuron drop.
Staphylococcus aureus Colonization in Community-Dwelling People With Spinal Cord Dysfunction

Mary-Claire Roghmann, MD, MS, Peter H. Gorman, MD, Mitchell T. Wallin, MD, Kristen Kreisel, MS, Simone Shurland, MS, Judith A. Johnson, PhD


Objectives: To estimate the prevalence of and determine risk factors for Staphylococcus aureus colonization of the perineum.

Design: Cross-sectional study with follow-up of up to 1 year.

Setting: Multiple outpatient sites.

Participants: Eighty-four community-dwelling adults with spinal cord dysfunction (SCD).

Interventions: Not applicable.

Main Outcome Measure: Colonization of perineum with S. aureus.

Results: Overall, 24% of the study cohort carried S. aureus on their perineal skin at enrollment, with 16% having methicillin-susceptible S. aureus and 10% having methicillin-resistant S. aureus (MRSA). Most perineal carriers were also colonized in the anterior nares. Participants with trauma as the cause of their SCD were more likely to be colonized with S. aureus than participants with SCD caused by multiple sclerosis or other causes (relative risk [RR], 2.8; 95% confidence interval [CI], 1.2–6.6; P=.01). Participants with pelvic decubiti were more likely to be colonized with S. aureus than participants without pelvic decubiti (RR=4.3; 95% CI, 2.4–7.7; P<.001). The recent use of any antibiotic was not associated with an increased risk of colonization with S. aureus (RR=1.5; 95% CI, 0.7–3.3; P=.31); however, recent fluoroquinolone use was significantly associated with perineal colonization (RR=2.8; 95% CI, 1.4–5.8; P=.02). Of the 8 participants with MRSA colonization, only 2 (25%) had a history of MRSA colonization.

Conclusions: S. aureus colonization of the perineum is common in this outpatient population of people with SCD. The use of fluoroquinolones was associated with S. aureus colonization. Colonization with MRSA without a history of MRSA colonization was common.

Key Words: Cross-sectional studies; Fluoroquinolones; Multiple sclerosis; Rehabilitation; Spinal cord injuries; Staphylococcus aureus.

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

STAPHYLOCOCCUS AUREUS IS a major cause of both community- and hospital-acquired infections.1 S. aureus infection is common in people with spinal cord injury (SCI) or dysfunction (SCD).2 Despite this, little is known about one of the major risk factors for S. aureus infection, colonization with S. aureus. Colonization with S. aureus is clearly important because rates of infection are higher in carriers than noncarriers.3 Studies comparing colonizing and infecting isolates of S. aureus have shown that people are usually infected with their own colonizing isolate.4 Temporary eradication of S. aureus by using topical mupirocin reduces S. aureus infection in certain populations.5,6 Understanding risk factors for S. aureus colonization can help us identify potential intervention to decrease S. aureus infection.7 Perineal colonization with S. aureus is common, affecting about 20% of the general population.8 Perineal carriers have high levels of colonization based on quantitative cultures.8 In addition, people with perineal or perirectal colonization may be at greater risk of S. aureus infection.9 S. aureus colonization of the nares, perineum, and wounds is common in people with SCI or SCD. Because of frequent hospitalization, these people are often colonized with methicillin-resistant S. aureus (MRSA), which is most commonly transmitted in the health care setting.10-12 Relatively little is known about S. aureus colonization in community-dwelling people with SCD. Our objectives in this investigation are to estimate the prevalence of S. aureus colonization of the perineum in a cross-sectional study of community-dwelling people with SCD and determine risk factors for perineal colonization. We also describe the patterns of colonization in a longitudinal study.

METHODS

Study Design

We prospectively enrolled community-dwelling adults with SCI or SCD (hereafter referred to collectively as SCD) from April 2003 to June 2004. The perineum of participants was cultured for S. aureus. We then assessed risk factors for S. aureus colonization in a cross-sectional study. We repeated their perineal cultures quarterly for up to 1 year to determine if colonization was persistent. We added a culture of the anterior nares to the protocol after the first 38 participants were enrolled when it became apparent from perineal culture results that S. aureus carriage was common. This assessment of S. aureus colonization was part of a larger review of antibiotic-resistance colonization patterns in this patient population, part of which...
was already published.\textsuperscript{13} Eligibility criteria and recruitment sites are described in detail in the prior publication.\textsuperscript{13,14} Participants provided informed consent and were followed longitudinally for up to 1 year. The institutional review boards of the University of Maryland, Baltimore, MD, and the Washington, DC, Veterans Affairs Medical Centers approved the study protocol. Implementation of this study conformed to the approved protocol, as well as the ethical and humane principles of research.

**Study Variables**

We collected age, sex, ethnic background, etiology of SCD, years since onset of SCD, level of disability as measured by the Eastern Cooperative Oncology Group (ECOG),\textsuperscript{15} presence of decubiti, method of emptying bladder, and previous hospitalizations and antibiotics within the last 90 days at enrollment from participants and their medical records. We classified participants with multiple sclerosis (MS) by disease subtype.\textsuperscript{16} We classified participants with SCI caused by trauma by their American Spinal Injury Association (ASIA) Impairment Scale.\textsuperscript{17} We categorized the 6-point ECOG score as 0 to 2 and 3 to 5 because people with an ECOG score of 0 to 2 are capable of all self-care, whereas people with an ECOG score of 3 to 5 are capable of only limited self-care or completely disabled.

We obtained cultures of the perineal skin on enrollment and then quarterly. Cultures were obtained by a single research coordinator. Perineal cultures were obtained after positioning the participant on his/her back with open legs by using a moistened, single-sterile Dacron fiber-tipped swab, inserting it into 1 anterior naris, rotating it around, and then repeating in the other nares.

Swabs were cultured for the presence of *S. aureus* including MRSA by using standard microbiologic methods. Swabs were inoculated on T-soy agar plates with 5% sheep blood agar, phenylethyl alcohol agar,\textsuperscript{4} and mannitol salt agar with and without 6μg/mL of oxacillin. Plates were incubated at 35°C and examined at 1- and 2-day growth. *S. aureus* was identified with BactiStaph latex agglutination.\textsuperscript{18} Methicillin resistance was confirmed on Mueller Hinton agar with 4% NaCl and 6μg/mL of oxacillin.\textsuperscript{18} Cultures negative for *S. aureus* on direct plating were tested via enrichment broth culture by using 2.5% NaCl broth\textsuperscript{18} and then direct plated as described earlier.

Persistent carriage was defined as greater than or equal to 80% of at least 5 perineal cultures positive for *S. aureus*; intermittent carriage was defined as 1% to 79% of at least 5 cultures positive for *S. aureus*, and no carriage was defined as 0% of at least 5 cultures. This definition is adapted from definitions used by VandenBergh et al\textsuperscript{19} to define patterns of carriage. Standard deoxyribonucleic acid (DNA) amplification and sequencing with the “universal” 27F and 519R primers of 16s recombinant DNA was performed to confirm all isolates as *S. aureus*.\textsuperscript{20} Each isolate was typed by DNA sequencing analysis of the protein A (*spa*) gene hyper-variable region as previously described.\textsuperscript{21}

**Study Analysis**

Summary statistics included means and standard deviations (SDs) or continuous variables and proportions for categorical variables. Group means were compared by using a *t* test or 1-way analysis of variance and proportions by the Pearson chi-square test or Fisher exact test, as appropriate. All statistical analyses were performed by using SPSS statistical software.\textsuperscript{8}

**RESULTS**

Table 1 shows a description of our participants. Forty-four participants (52%) had MS, and 36 (45%) had trauma as a cause of their SCD. Of those with MS, 45% were classified as relapsing-remitting and 43% as secondary progressive. The remaining were classified as primary progressive (n = 3) or not specified (n = 2). Of those with SCI caused by trauma, 58% were ASIA grade A, 11% were grade B, 19% were grade C, and 11% were grade D. Of the remaining participants, 1 had Brown-Séquard syndrome, 2 had central cord syndrome, and 1 had cauda equina syndrome. Twenty-six percent had used antibiotics in the 90 days before study enrollment, and 12% had used fluoroquinolone antibiotics. Six percent had a history of MRSA colonization. Overall, 24% of the study cohort carried *S. aureus* on their perineal skin at enrollment, with 16% having methicillin-sensitive *S. aureus* (MSSA) and 10% having MRSA. One participant had both MRSA and MSSA colonization. Of the 8 participants with MRSA colonization, only 2 (14%) had a history of MRSA colonization. All but one of the MRSA isolates had *spa* motifs (MDMGMK) associated with the hospital-acquired MRSA clones USA100 or USA800.\textsuperscript{22}

Table 1 also shows characteristics associated with *S. aureus* colonization of the perineal skin. Participants with trauma as the cause of their SCD were more likely to be colonized with *S. aureus* than participants with SCD caused by MS or other causes (relative risk [RR], 2.8; 95% confidence interval [CI], 1.2–6.6; *P* = .01). Participants with pelvic decubiti were more likely to be colonized with *S. aureus* than participants without pelvic decubiti (RR = 4.3; 95% CI, 2.4–7.7; *P* < .001). The recent use of any antibiotic was not associated with an increased risk of colonization with *S. aureus* (RR = 1.5; 95% CI, 0.7–3.3; *P* = .31); however, recent fluoroquinolone use was significantly associated with perineal colonization (RR = 2.8; 95% CI, 1.4–5.8; *P* = .02). Among 38 participants with trauma as the cause of their SCI, a greater proportion of those who had taken fluoroquinolones in the past 90 days were *S. aureus* colonized compared with those who had not (60% vs 29%, *P* = .026). Most perineal carriers were also colonized in the anterior nares. Of the 46 participants who were cultured in the anterior nares, 13 of 16 perineal carriers were also nasal carriers, and only 4 of the 30 nonperineal carriers were nasal carriers (81% vs 13%, *P* = .0001).

Patterns of *S. aureus* carriage were assessed in 30 participants who were followed for 1 year with at least 5 perineal cultures. The demographics, characteristics of SCD, and potential risk factors for *S. aureus* colonization of these participants did not differ from the participants who dropped out (data not shown). Overall, 10% were persistently colonized with 80% or more of perineal cultures positive, 63% were intermittently colonized with 20% to 60% of cultures positive, and 27% were never positive for *S. aureus*. Among 23 participants with *S. aureus* colonization, we detected 22 different *spa* types. Three of the 19 intermittent carriers showed evidence of a switch in *spa* type, whereas 1 of the 3 persistent carriers had a modification in *spa* type (16% vs 33%, *P* = 1.0). We had cultures of the anterior nares on 22 of the 30 participants in the longitudinal component of the study. Among perineal carriers of *S. aureus* with a nares culture, 11 of 16 were nasal carriers of *S. aureus*; among nonperineal carriers of *S. aureus* with a nares culture, 1 of 6 were nasal carriers. Among the 11 with
nasal and perineal carriage, all but I carried the same *spa* type in the anterior nares as in the perineum.

**DISCUSSION**

**Summary of Main Results Placed in Context of Other Studies**

Overall, 24% of the study population carried *S. aureus* on their perineal skin at enrollment. People without SCD are also known to carry *S. aureus* in their perineum with similar frequency. In 1959, Ridley* cultured 50 male medical students. He found that overall 48% were nasal and perineal carriers, all but 1 carried the same *S. aureus* type.

Other Studies

<table>
<thead>
<tr>
<th>Variables</th>
<th>Overall Population (N=84)</th>
<th>Perineal Colonization With <em>S. aureus</em> at Enrollment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male, n (%)</td>
<td>70 (83)</td>
<td>16 (80)</td>
</tr>
<tr>
<td>Mean age ± SD (y)</td>
<td>50±10</td>
<td>48±10</td>
</tr>
<tr>
<td>Ethnic background</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White, non-Hispanic, n (%)</td>
<td>55 (65)</td>
<td>11 (55)</td>
</tr>
<tr>
<td>African American, non-Hispanic, n (%)</td>
<td>27 (32)</td>
<td>8 (40)</td>
</tr>
<tr>
<td>Other, n (%)</td>
<td>2 (2)</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Characteristics of SCD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Etiology of SCD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trauma, n (%)</td>
<td>38 (45)</td>
<td>14 (70)</td>
</tr>
<tr>
<td>MS, n (%)</td>
<td>44 (52)</td>
<td>6 (30)</td>
</tr>
<tr>
<td>Other, n (%)</td>
<td>2 (2)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Mean years with SCD ± SD</td>
<td>17±12</td>
<td>14±10</td>
</tr>
<tr>
<td>ECOG score of 3-5, n (%)</td>
<td>14 (17)</td>
<td>4 (20)</td>
</tr>
<tr>
<td>Presence of skin breakdown in pelvic area, n (%)</td>
<td>16 (19)</td>
<td>8 (40)</td>
</tr>
<tr>
<td>Method of emptying bladder</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continent, n (%)</td>
<td>42 (50)</td>
<td>6 (30)</td>
</tr>
<tr>
<td>Condom catheter, n (%)</td>
<td>12 (14)</td>
<td>2 (10)</td>
</tr>
<tr>
<td>Intermittent catheterization, n (%)</td>
<td>15 (18)</td>
<td>4 (20)</td>
</tr>
<tr>
<td>Suprapubic catheter, n (%)</td>
<td>3 (4)</td>
<td>3 (15)</td>
</tr>
<tr>
<td>Urethral catheter, n (%)</td>
<td>6 (7)</td>
<td>3 (15)</td>
</tr>
<tr>
<td>Incontinent (no catheter use), n (%)</td>
<td>6 (7)</td>
<td>2 (10)</td>
</tr>
<tr>
<td>Potential risk factors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>History of MRSA colonization, n (%)</td>
<td>5 (6)</td>
<td>3 (16)</td>
</tr>
<tr>
<td>Hospitalizations &gt;48h in last 90d, n (%)</td>
<td>5 (6)</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Antibiotics in last 90d, n (%)</td>
<td>22 (26)</td>
<td>7 (35)</td>
</tr>
<tr>
<td>Fluoroquinolone antibiotics in last 90d, n (%)</td>
<td>11 (13)</td>
<td>6 (30)</td>
</tr>
</tbody>
</table>

*P* values were calculated using Fisher’s exact test or Chi-square test as appropriate. Where *P* < .05, the results are considered statistically significant.

Arch Phys Med Rehabil Vol 88, August 2007

---

Table 1: Overall Description of Community-Dwelling People With SCD and Characteristics Associated With Perineal Colonization With *S. aureus* at Enrollment

<table>
<thead>
<tr>
<th>Variables</th>
<th>Overall Population (N=84)</th>
<th>Perineal Colonization With <em>S. aureus</em> at Enrollment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male, n (%)</td>
<td>70 (83)</td>
<td>16 (80)</td>
</tr>
<tr>
<td>Mean age ± SD (y)</td>
<td>50±10</td>
<td>48±10</td>
</tr>
<tr>
<td>Ethnic background</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White, non-Hispanic, n (%)</td>
<td>55 (65)</td>
<td>11 (55)</td>
</tr>
<tr>
<td>African American, non-Hispanic, n (%)</td>
<td>27 (32)</td>
<td>8 (40)</td>
</tr>
<tr>
<td>Other, n (%)</td>
<td>2 (2)</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Characteristics of SCD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Etiology of SCD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trauma, n (%)</td>
<td>38 (45)</td>
<td>14 (70)</td>
</tr>
<tr>
<td>MS, n (%)</td>
<td>44 (52)</td>
<td>6 (30)</td>
</tr>
<tr>
<td>Other, n (%)</td>
<td>2 (2)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Mean years with SCD ± SD</td>
<td>17±12</td>
<td>14±10</td>
</tr>
<tr>
<td>ECOG score of 3-5, n (%)</td>
<td>14 (17)</td>
<td>4 (20)</td>
</tr>
<tr>
<td>Presence of skin breakdown in pelvic area, n (%)</td>
<td>16 (19)</td>
<td>8 (40)</td>
</tr>
<tr>
<td>Method of emptying bladder</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continent, n (%)</td>
<td>42 (50)</td>
<td>6 (30)</td>
</tr>
<tr>
<td>Condom catheter, n (%)</td>
<td>12 (14)</td>
<td>2 (10)</td>
</tr>
<tr>
<td>Intermittent catheterization, n (%)</td>
<td>15 (18)</td>
<td>4 (20)</td>
</tr>
<tr>
<td>Suprapubic catheter, n (%)</td>
<td>3 (4)</td>
<td>3 (15)</td>
</tr>
<tr>
<td>Urethral catheter, n (%)</td>
<td>6 (7)</td>
<td>3 (15)</td>
</tr>
<tr>
<td>Incontinent (no catheter use), n (%)</td>
<td>6 (7)</td>
<td>2 (10)</td>
</tr>
<tr>
<td>Potential risk factors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>History of MRSA colonization, n (%)</td>
<td>5 (6)</td>
<td>3 (16)</td>
</tr>
<tr>
<td>Hospitalizations &gt;48h in last 90d, n (%)</td>
<td>5 (6)</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Antibiotics in last 90d, n (%)</td>
<td>22 (26)</td>
<td>7 (35)</td>
</tr>
<tr>
<td>Fluoroquinolone antibiotics in last 90d, n (%)</td>
<td>11 (13)</td>
<td>6 (30)</td>
</tr>
</tbody>
</table>

*P* values were calculated using Fisher’s exact test or Chi-square test as appropriate. Where *P* < .05, the results are considered statistically significant.
is consistent with what has been noted in other populations when assessing rates of nasal carriage. Previous work in other populations have found that patients with intermittent carriage are more likely to have different *S. aureus* strains over time, implying that these people are becoming colonized, then losing their isolates, and then acquiring a new isolate. The majority of our participants with intermittent carriage were colonized with a single *spa* type of *S. aureus* despite intermittent carriage. This implies that in this population, the intermittent carriers are truly persistently colonized with a single strain and that negative cultures are false-negatives.

In those participants for whom we had both perineal and nare cultures, we found that perineal cultures were strongly associated with nasal cultures. Eighty-one percent of those with positive perineal cultures had positive nare culture, and 76% of those with positive nare cultures had positive perineal cultures. This is in contrast to other studies that found that the anterior nares was more commonly colonized than the perineum; our use of enrichment broth cultures may explain the difference. Alternatively, perineal colonization may be more common in this population. This strong correlation suggests that our results may apply to *S. aureus* colonization at either site.

**Study Strengths**

Our study has a number of important strengths. We studied a unique population of people with SCD in a prospective manner, thereby optimizing the accuracy of our data. Few, if any, studies have looked at *S. aureus* colonization in a community-dwelling population of people with SCI; most studies have been performed in the inpatient setting. We also used surveillance cultures, not clinical cultures (which are taken in response to the clinical suspicion of an infection), to measure the presence of *S. aureus*, and our culture methods included the use of enrichment broth for the detection of *S. aureus* maximizing the sensitivity of the cultures to detect *S. aureus*.

**Study Limitations**

Our cross-sectional study was limited by a relatively small sample size. Because of this, we cannot say whether the use of fluoroquinolones is an independent risk factor for *S. aureus* colonization. The longitudinal study was limited because of loss to follow-up; however, those lost to follow-up were similar to those who participated for a full year. Thus, our estimates of the distribution of *S. aureus* carriage over time should be accurate for our target population.

**CONCLUSIONS**

*S. aureus* colonization of the perineum is common in this outpatient population of people with SCD; if colonized, most carry a single strain for over a year. The use of fluoroquinolones may promote *S. aureus* colonization, which is associated with an increased risk of *S. aureus* infection. Thus, we need to use fluoroquinolones to treat only symptomatic infections, not asymptomatic colonization or contamination.

**Acknowledgments:** We thank Debbie Grady, RN, for her critical role in coordinating the project. We thank the clinical staff of enrollment sites and other members of the research team including the VA MS Center of Excellence-East. This assessment of *S. aureus* colonization was part of a larger review of antibiotic resistance colonization patterns in this patient population, part of which was previously published.

**References**


Suppliers
a. Remel Inc, 12076 Santa Fe Dr, PO Box 14428, Lenexa, KS 66215.
b. Version 12.1; SPSS Inc, 233 S Wacker Dr, 11th Fl, Chicago, IL 60606.
Role of Sonographic Examination in Traumatic Knee Internal Derangement

Chung-Yuan Wang, MD, Hsing-Kuo Wang, PhD, Chao-Yu Hsu, MD, Jeng-Yi Shieh, MD, Tyng-Guey Wang, MD, Ching-Chuan Jiang, PhD


Objectives: To define the accuracy (compared with magnetic resonance imaging [MRI]) of sonographic examination in detecting knee effusion and to determine whether the presence of knee effusions in patients with traumatic knee injury can predict knee internal derangement as assessed by MRI.

Design: Prospective study.

Setting: Hospital rehabilitation department.

Participants: Thirty patients (19 men, 11 women) with traumatic knee injury were recruited. Subjects received sonographic examination and MRI on the same day.

Interventions: Not applicable.

Main Outcome Measures: The presence or absence of knee effusion was assessed by sonographic examination. MRI was used as criterion standard to evaluate whether the presence of knee effusion and internal derangement, which included tear of anterior and posterior cruciate ligaments, as well as meniscus tear.

Results: The sensitivity of sonographic examination for detecting knee effusion was 79.1%, and specificity was 50%. The positive-predictive value (PPV) was 86.3% and negative-predictive value (NPV) was 37.5%. The PPV of sonographic effusion to internal derangement was 90.9%, and the NPV was 37.5%.

Conclusions: Sonographic examination can accurately detect effusion of the knee. The detection of knee effusion in patients with traumatic knee injury by sonographic examination is highly indicative of internal knee derangement.

Key Words: Doppler ultrasound; Knee injuries; Magnetic resonance imaging; Rehabilitation.

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

Sonographic examination has been widely used for detecting musculoskeletal disorders in recent decades.1,2 It has the advantages of being noninvasive, readily available, dynamic, and allows for good visualization of superfluous structures.3,4 It is well accepted for use in evaluating extra-articular structures of the knee,4,6 but is still debated in intra-articular examination.5 Although sonographic examination is not perfect for evaluating the internal structure of the knee, it is thought to be accurate for the detection of knee effusion.3,7,8 However, the clinical role of sonographic examination in detection of knee effusion is not well explored.

Knee internal derangement is common in traumatic knee injury. Patient history, physical examination, imaging studies, and arthroscopy all play important roles in the diagnosis of internal derangement of knee.9 Of all, arthroscopy is considered to be the criterion standard for assessing knee internal derangement,3,10 but it has the downside of being an invasive procedure. Magnetic resonance imaging (MRI) is highly accurate for detecting intra-articular lesions of the knee,3 and is often used to assess patients with suspected internal knee disorders as an alternative to arthroscopy.11,12 Esmaili Jah et al13 have suggested that skilled physical examination may be more accurate than MRI for detecting anterior cruciate ligament (ACL), posterior cruciate ligament (PCL), and meniscal injuries. However, the quality and accuracy of physical examination are highly dependent on examiner skill and experience. Therefore, MRI is generally recommended on patients with acute knee injury and equivocal physical findings.9 Nevertheless, MRI has the disadvantages of high cost and time required for scheduling. Having the advantages of time-efficiency, economic, and acceptable accuracy in detecting musculoskeletal disorders, sonography is considered to be an alternative tool.

Knee effusion is a common sign of knee pathology.13,14 It may result from trauma, overuse, or systemic disease.14 The most common underlying causes of traumatic knee effusion are due to ligamentous, meniscal, or osseous injury.14 Several studies have shown correlation between knee effusion and internal derangement,15,16 but the patients examined were either markedly divergent in age, or had widely varying criteria of internal derangement. Moreover, there is no study using the ultrasound to assess the presence of knee effusion, and further to predict whether knee internal derangement existed. Hence the aim of our study was to define the accuracy of sonographic examination for detection of knee effusion, and furthermore to determine whether the presence of knee effusion in patients with traumatic knee injury is predictive of knee internal derangement.

Methods

Participants

For this work, we recruited 30 consecutive patients (19 men, 11 women) with traumatic knee injury referred for MRI study. The mean age of the subjects was 27 years (range, 16–52y).

Knee injuries were associated with the following traumas: traffic accident (n=8), sports injury (n=10), and falls while performing various other activities (n=12). Patients were examined by sonography and MRI on the same day by 2 blinded, independent investigators. The study was approved by the ethics committee of the university hospital and all subjects provided informed consent prior to participating in the work.
spin-echo T2-weighted images. Slice thickness was 4mm for T2-weighted, sagittal spin-echo T1-weighted, and sagittal fast was performed as follows; coronal multiplanar gradient recall commercially available circumferential extremity coil. Imaging transducer in a portable ultrasound was used to scan knee loskeletal sonographic experience). A 10-MHz linear array the author group, an expert in sonography (5 years of muscu- view.

Protocol
Subjects received ultrasound examination before the MRI study. The sonographic examination focused on the detection of knee effusion. MRI was used to define both internal derangement and effusion. Internal derangements in this study were defined as ACL, PCL, or meniscal lesion. The clinical history and physical examination were recorded by chart re- view.

Sonography
Sonographic examinations were performed by a member of the author group, an expert in sonography (5 years of muscu-
loskeletal sonographic experience). A 10-MHz linear array transducer in a portable ultrasound was used to scan knee joints. Subjects were maintained in a supine position with knee full extension, and the transducer was aligned along the longi-
tudinal axis of the knee joint, with the lower end of transducer at the upper margin of patella. In this position, the suprapatellar recess was located between the quadriceps tendon and prefemoral fat. Then the transducer was swept to the medial and lateral site of quadriceps tendon to fully scan the suprapa-
tellar recess. Effusions appeared as an anechoic area between the quadriceps tendon and prefemoral fat (fig 1). In patients with mild effusion, it could be only detected in the lateral or medial compartment due to effect of gravity. The effusion was recorded as the widest anechoic width on sonogram without compressing the transducer or quadriceps contraction. Sonogram criteria for the presence of knee effusion was fluid accumulation exceeding 2mm, because the physiologic fluid accumulation is usually less than 2mm. To be certain that anechoic area was effusion rather than synovium or other soft tissue mass, Doppler examination and compressing the area were performed. Knee joint fluid was highly compressible, and did not exhibit vascularity.

Magnetic Resonance Imaging
MRI studies were performed using a 1.5-T magnet and a commercially available circumferential extremity coil. Imaging was performed as follows; coronal multiplanar gradient recall T2-weighted, sagittal spin-echo T1-weighted, and sagittal fast spin-echo T2-weighted images. Slice thickness was 4mm for sagittal and coronal images. The interslice gap was 1mm. The matrix size was 256×192, and the field of view was 14 to 16cm.

We evaluated MR images for the presence of knee effusions and internal derangement. MRI criteria for the presence of knee effusion were fluid accumulation in the suprapatellar recess exceeding 10mm in anteroposterior width on sagittal images. MRI indicators of ACL tear included: (1) an irregular, wavy contour to the anterior margin of the ACL, (2) high signal intensity within the substance of the ACL on T2-weighted images, (3) discontinuity of the substance of the ACL on sagittal images, and (4) secondary signs such as buckling of the PCL and anterior subluxation of the tibia. MRI indicators of complete PCL tear included: (1) failure to identify the PCL, (2) amorphous high signal intensity in the region of the PCL on T1- and T2-weighted images without definable ligamentous fibers, and (3) visualization of PCL fibers with focal discrete disruption of all visible fibers. Partial PCL tear or intrasub-

NOTE. Sensitivity, 79.1%; specificity, 50%; positive-predictive value, 86.3%; and negative-predictive value, 37.5%.

Table 1: Results of Sonographic Examination With Respect to MRI Diagnosis (N=30)

<table>
<thead>
<tr>
<th>Sonographic Diagnosis</th>
<th>MRI Diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effusion (positive)</td>
<td>19</td>
</tr>
<tr>
<td>Effusion (negative)</td>
<td>5</td>
</tr>
</tbody>
</table>

NOTE. Sensitivity, 80%; specificity, 60%; PPV, 90.9%; and NPV, 37.5%.

Table 2: The Relationship Between Knee Sonographic Effusion and Internal Derangement in MRI (N=30)

<table>
<thead>
<tr>
<th>Sonographic Effusion</th>
<th>Internal Derangement in MRI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>20</td>
</tr>
<tr>
<td>Negative</td>
<td>5</td>
</tr>
</tbody>
</table>

NOTE. Sensitivity, 80%; specificity, 60%; PPV, 90.9%; and NPV, 37.5%.
was 80.0% and the specificity was 60.0%. The PPV and NPV were 90.9% and 37.5%, respectively.

Of the 25 patients with knee internal derangement, 10 had multiple injuries. Fifteen had ACL tears, 5 PCL tears, and 15 meniscus tears in these 25 patients. There were 5 patients who had internal derangement but no evident knee effusion on sonogram, whereas 2 patients had sonographic knee effusion in the absence of internal derangement. Of the 5 patients exhibiting internal derangement but no sonographic effusion, 3 had no effusion in MRI either, and the interval between their injuries was almost 3 months. The other 2 subjects had only a little effusion in MRI and the interval between their injuries was 5 days and 1 month. These 5 subjects were individually diagnosed with the following: one had ACL and meniscal injury, one ACL injury alone, and 3 meniscal injury. Of the 2 patients with sonographic effusion but no internal derangement in MRI, one had a mild synovitis, and one had a severe inflammatory arthritis with synovitis.

**DISCUSSION**

Findings from this study indicate that sonographic examination is a useful imaging tool for detecting knee effusion, with an accuracy of around 73%. There were 5 patients in whom effusion was detected by MRI, but not by sonographic analysis. The extent of effusion detected in all of these particular patients by MRI was minimum amount, suggesting that sonographic examination may be sufficiently accurate to detect knee effusion in all but the most minimal cases. In the study by Schweitzer et al.19 MRI could detect as little as 4mL of fluid at mid-sagittal line. However, Delaunois et al7 stated that the lowest amount of effusion detected by sonography in saline or blood was 10mL. These findings may imply that sonographic examination is less sensitive than MRI in detecting knee effusion. Whether such amounts of knee effusion have clinical significance still needs to be determined.

Knee effusion can also be detected by physical examination and can be used to predict the knee internal derangement. Good validity of physical examination in detecting a large amount of knee effusion was reported by Hauzeur et al20 but poor reliability was noted when the effusion was moderate to small. Obese, joint deformities in osteoarthritis, and edema due to venous disease might further increase the clinical evaluation errors.21 Differentiating joint effusion from hypertrophy of the synovium is sometimes difficult, either by physical examination or ultrasound. However, the compression maneuver on sonographic examination made it easy to differentiate synovium from effusion, and the compression maneuver was performed in this work. Hauzeur2 classified the amount of knee effusion into 3 grades on sonogram. Large effusion was defined as fluid collection evident in the suprapatellar pouch, moderate effusion as fluid detected only when the lateral pouches were being compressed, no effusion as no fluid collection even by compression. We did not classify the amount of effusion in our subjects due to the small number of cases. It might be interesting to classify the effusion amount in the further study.

The other important finding of this study is that subjects with knee effusion detected by ultrasound had a 90.9% PPV of knee internal derangement. This suggests that when knee effusion is detected by sonographic examination, it is highly likely that knee internal derangement also exists, and that further diagnostic evaluation (by MRI or arthroscopy) is warranted. Sonographic examination could hence be used as a valuable screening tool for patients with traumatic knee injury, preventing the need for the more costly and invasive techniques of MRI and arthroscopy, respectively.

In 1 retrospective study,9 59% of 115 patients with knee effusion (as diagnosed by MRI) had internal derangement. The correlation between knee effusion and internal derangement was significantly lower than the present study. However, in another study,20 90% of acute knee injury patients were found to have both effusion and internal derangement as determined by MRI. Several factors may influence the correlation between the knee effusion and internal derangement. First, the duration between injury and evaluation should be considered, because effusions absorb over time. This was shown by Boks et al,21 who noted that effusion was related to recent, but not old, trauma. In our study, most of the patients visited the clinic several days or even weeks after the knee injury occurred. Second, the amount or extent of effusion is a possible determining factor. Small amounts of effusion may result from inflammation of the knee joint, rather than injury of the knee structure. For instance, in a study conducted by Kolman et al,13 it was determined that only 14% of patients with 10mm or less of effusion in the lateral aspect of the suprapatellar pouch had concurrent derangement. It was concluded that there is no significant effusion evident in the suprapatellar pouch there is unlikely to be associated derangement. Last, the differing definitions of internal derangement may influence the results. Some researchers have included bony fracture or dislocation as internal derangement, but we did not. Internal derangement was broadly defined in the study conducted by Kolman,13 with anterior, posterior, median, and lateral cruciate ligament tears, meniscal tears, moderate to severe osteoarthritis, osteochondrosis dissecans lesions, patellar microfracture, bone bruising, tibial plateau fracture, tibial epiphyseal fracture, patellar retinacular tears, and popliteal tears all included. In keeping with most of the studies,5,11 we defined internal derangement as by the presence of either ACL, PCL, or meniscal injury.

**Study Limitations**

Selection bias did exist in our study. The patients who were referred to performing MRI study represent a group highly suspected to have internal derangement clinically. This bias is apparent in the finding that only 5 patients in this work did not have internal derangement. Selection bias was also apparent with regard to the injury distribution of internal derangement. In the study by Luhmann,22 the following distribution was observed: 29% with ACL injuries, 29% with meniscal tears, 25% with patellar retinacular tears, patellar microfracture, bone bruising, tibial plateau fracture, tibial epiphyseal fracture, patellar retinacular tears, and popliteal tears all included. In our study population, the distribution was similar to the previous reports. Court-Payen3 had suggested that sonographic examination is useful for the diagnosis of 4 conditions: inflammatory joint diseases, peri-articular masses, suspected meniscus or ligament lesions, and loose bodies. In addition to this, Khan et al3 concluded that sonographic examination is accurate in evaluating knee internal derangement. However, sonographic accuracy in intra-articular examination was still debated, though the resolution of sonography had improved greatly. Azzoni and Cabitza23 found that the accuracy of sonographic examination for detecting meniscus tear was as low as 44%. The major reasons for nondiagnosis were inability to view the entire meniscus on the sonogram and the presence of artifacts on adjacent bone surfaces. Regarding the ACL, there has been no report detailing direct visualization of it by sonography, although some authors have attempted indirect evaluation.8,24,25
Several reports have described that torn PCL has increased thickness, reduced echogenicity, and focal disruption when viewed sonographically. The sensitivity and specificity in these studies were not discussed. The accuracy of ultrasound in assessing meniscus and internal ligament lesions are not well established but this study suggests a possible role of sonographic examination in traumatic knee injury. Sonographic examination can accurately detect the presence of a knee effusion, and help determine the necessity of further image or invasive study in this group of patients.

CONCLUSIONS

Sonographic examination can accurately detect knee effusion. When knee effusion is evident in patients of traumatic knee injury, it is strongly indicative of knee internal derangement. We suggest that sonography is an ideal screening tool to determine whether more detailed knee examination (ie, to diagnose internal derangement) is warranted.

Acknowledgment: We thank Luke Carey, PhD, of Medica Communionis Asia, for his valuable comments and editorial services.

References


Suppliers

a. SonoSite Inc, 21919 30th Dr SE, Bothell, WA 98021.
b. Sigma; GE Healthcare USA Inc, 8200 W Tower Ave, Milwaukee, WI 53223.
c. Siemens Medical Solutions USA Inc, 51 Valley Stream Pkwy, Malverne, NY 11565.

Objectives: To describe the prevalence of greater trochanteric pain syndrome (GTPS); to determine whether GTPS is associated with ilioitibial band (ITB) tenderness, knee osteoarthritis (OA), body mass index (BMI), or low back pain (LBP); and to assess whether GTPS is associated with reduced hip internal rotation, physical activity, and mobility.

Design: Cross-sectional, population-based study.

Setting: Multicenter observational study.

Participants: Community-dwelling adults (N=3026) ages 50 to 79 years.

Interventions: Not applicable.

Main Outcome Measures: Greater trochanter tenderness to palpation in subjects with complaints of hip pain and no signs of hip OA or generalized myofascial tenderness.

Results: The prevalence of unilateral and bilateral GTPS was 15.0% and 8.5% in women and 6.6% and 1.9% men. Odds ratio (OR) for women was 2.79 (95% confidence interval [CI], 2.27-4.25), but age and race were not significantly associated with GTPS. In a multivariate model, adjusting for age, sex, ITB tenderness, ipsilateral and contralateral knee OA, BMI, and LBP, ITB tenderness (OR=1.72; 95% CI, 1.34-2.19), knee OA ipsilaterally (OR=3.47; 95% CI, 2.72-4.42) and contralaterally (OR=1.74; 95% CI, 1.32-2.28), and LBP (OR=2.79; 95% CI, 2.22-3.50) were positively related to GTPS. In this complete model, BMI was not associated with GTPS (OR=1.10; 95% CI, 0.80-1.52 when comparing 30 with <25kg/m²). Hip internal rotation range of motion did not differ based on GTPS status. After multivariate adjustment, GTPS did not alter physical activity score, but bilateral GTPS was significantly associated with a higher 20-meter walk time and chair stand time.

Conclusions: The higher prevalence of GTPS in women and in adults with ITB pain or knee OA indicates that altered lower-limb biomechanics may be related to GTPS. Slower functional performance in those with GTPS suggests that the study of targeted rehabilitation may be useful. A longitudinal study will be necessary to identify causal factors and outcomes of interventions.

Key Words: Bursitis; Femur; Rehabilitation.

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

Greater Trochanteric Pain Syndrome: Epidemiology and Associated Factors

Neil A. Segal, MD, David T. Felson, MD, James C. Torner, PhD, Yanyan Zhu, MSc, Jeffrey R. Curtis, MD, MPH, Jingbo Niu, DSc, Michael C. Nevitt, PhD, for the Multicenter Osteoarthritis (MOST) Study Group

GREATER TROCHANTERIC pain syndrome (GTPS) is defined as tenderness to palpation over the greater trochanter with the patient in the side-lying position. In contrast to the term greater trochanteric bursitis, which implies presence of inflammation, referring to this clinical entity as GTPS is preferable for 2 reasons: (1) pain in this region frequently is not associated with signs of inflammation such as warmth, erythema, or swelling, and (2) the etiology is not fully known and may relate to myofascial pain rather than inflammation. Pain generators may be associated with the gluteus maximus, medius, or minimus bursae; muscle attachments; or overlying tissue such as the ilioitibial band (ITB).

The myofascial attachments and bursae associated with the greater trochanter may be affected by altered lower-limb biomechanics. Physical medicine and rehabilitation and orthopedics textbooks cite that osteoarthritis (OA) of the lumbar spine, hip, or knee; ITB tightness or tendinitis; or strain of the hip external rotators may contribute to trochanteric pain by adding stress to the area. Other experts have written that trochanteric bursitis is more prevalent in women, frequently associated with mechanical by a strain and obesity and often with reduced hip internal rotation range of motion (ROM). However, others suggest that marked loss of hip internal rotation is typically not associated with trochanteric pain. One reason for the differences of opinion regarding these and other factors hypothesized to be associated with GTPS is that none have been analytically assessed.

The prevalence of GTPS in adults with musculoskeletal low back pain (LBP) has been reported to be 20% to 35%. Studies differ regarding whether GTPS may or may not be more prevalent in women than men. However, numerous sources have recognized that there is a lack of data regarding the prevalence of GTPS. One study suggested that GTPS was the cause of LBP, based on findings that Oswestry Disability Index scores for female patients with LBP improved after greater trochanteric steroid injection. However, considering that the corticosteroid is systemically absorbed, this uncontrolled study may not answer whether there is an association between back pain and GTPS. Although there have been small studies of GTPS in back pain clinics, we are unaware of any study examining the prevalence and epidemiology of GTPS in the general population. For providers of musculoskeletal care, it is important to...
understand factors associated with GTPS to minimize significant decline of physical function and quality of life.

The long delay between presentation and diagnosis of hip external rotator muscle tears, 1 cause of GTPS, was recently found to be caused by underrecognition by physicians. Such underrecognition may be because of a lack of education or attention. Frequently, GTPS is associated with pain at night, pain when standing greater than 15 minutes, or radiating pain and paresthesias. These complaints may lead to ordering tests such as magnetic resonance imaging or electromyography to assess for radiculopathy instead of conducting a well-informed physical examination.

Understanding whether GTPS is associated with sex, body mass index (BMI), or other lower-limb musculoskeletal diagnoses would enable future study of risk factors for incident and progressive GTPS. Additionally, the characterization of activity level or hip flexibility that could affect pathomechanics at other joints in the kinetic chain, such as the knee or spine, would also guide rehabilitation considerations.

To address these needs, this study assessed the prevalence of GTPS in a community-based population complaining of lower-limb pain and whether GTPS was associated with 4 purported risk factors: (1) ipsilateral ITB pain, (2) knee OA, (3) obesity, and (4) LBP. A secondary aim was to determine whether hip internal rotation ROM or levels of physical activity and physical performance are limited in adults with GTPS.

METHODS

Participants

The Multicenter Osteoarthritis (MOST) Study is a longitudinal study of persons 50 to 79 years old at a high risk of developing symptomatic knee OA or who already had knee OA. Subjects are drawn from the community. Although the purpose of the MOST study is to assess risk factors for incident and progressive knee OA, the availability of standardized data on GTPS diagnosis and associated features in a large community-based sample makes it a useful resource for analysis. The present report uses data from the baseline examination of the MOST study in which subjects underwent a musculoskeletal examination. Institutional review board approval was obtained at each of the investigators’ institutions before initiating recruitment and research protocols.

Subject recruitment was conducted by MOST study personnel who contacted adults aged 50 to 79 years by mass mailings as well as through printed advertisements in the counties surrounding the 2 clinical sites. Potential subjects who indicated interest were then contacted by telephone for screening. Subjects were recruited if they had a history of knee pain, injury, or surgery or were overweight. Exclusion criteria included bilateral knee replacement, cancer, or rheumatologic disease.

Demographic variables and a report of knee pain, hip pain, and LBP were assessed through questionnaires administered by telephone and in the clinic. At the baseline MOST visit, the subset of 1786 subjects who indicated “pain, aching, or stiffness” located on the lateral aspect of either hip on most days over the last month on a pain diagram underwent a physical examination. Subjects with a prosthetic hip were excluded from the physical examination. We used this physical examination to acquire information on tenderness over the greater trochanter. Subjects who indicated no lateral hip pain were considered not to have GTPS.

Physical Examination

GTPS was defined as tenderness on physical examination in the absence of generalized myofascial tenderness to palpation. Examiners used a Chatillon CMD 10-1 dolorimeter to calibrate finger pressure to 1.4 and 3.0kg of pressure before palpating subjects’ greater trochanters. Examiners asked subjects “is this tender or painful” while applying 1.4 to 3.0kg of thumb-tip pressure over the lateral and posterior aspects of each greater trochanter with the subject in the lateral decubitus position.” A positive response to this question was defined as tenderness in the greater trochanter region. Generalized myofascial tenderness was defined by an affirmative response to the same question when 1.4kg of pressure was applied over the soft tissue 2cm proximal to the medial joint line of the knee as well as at 2 or more of the following points: left and right proximal trapezius and left and right extensor mass immediately distal to the lateral epicondyle of an elbow.

ITB tenderness also was assessed and defined as an affirmative response to “is this tender or painful” when applying 1.4kg of pressure with the thumb over each ITB just proximal to the lateral femoral condyle with the subject in the lateral decubitus position. Internal rotation of each hip was measured from neutral, using a 25.4cm (10-in) long-arm goniometer, measuring to the nearest 1°, with subjects seated on the edge of the examination table with hips and knees flexed to 90°. All subjects were examined using a standardized, written protocol, with regular reliability checks performed to ensure consistency between sites and examiners.

LBP was defined as being “bothered by back pain most or all of the time” over the last 30 days. Subjects also were asked “During the past 30 days, have you limited your activities due to back pain?”

Radiograph protocol for the assessment of knee OA in the MOST study has been described previously. Knee OA was defined as at least 1 definite osteophyte visible at standard image size on posteroanterior knee radiographs. BMI was calculated as weight (in kilograms) per squared height, and obesity was defined as a BMI of 30kg/m² or higher. Subjects were encouraged to empty their bladders and bowels, empty their pockets, and remove jewelry before stepping on the scale with paper shorts and a shirt. Weight was measured with a standard medical beam balance with the certified examiner standing behind the subject and following a written protocol. Weight was recorded to the nearest 0.1kg immediately after the measurement. The scale was calibrated monthly with a 50-kg weight for accuracy as well as 5-, 10-, 15-, and 20-kg weights for linearity calibration. Additionally, the scale was calibrated annually by the local Department of Weights and Measures.

Height was measured with a wall-mounted Harpenden stadiometer and followed a written protocol. Subjects stood without shoes (barefoot or thin stockings) with their heels together and scapulae, buttocks, and both heels touching the wall plate with their head in the Frankfort horizontal plane. A standardized script and positioning protocol were used to measure height on full inspiration to the nearest 1mm. A 0.5-kg soft weight was placed on the headboard to standardize pressure on the head during measurement. The measurement was repeated twice, and, if these differed by more than 3mm, 2 additional measurements were taken. The stadiometer was calibrated daily with a 600-mm rod.

The activity level was measured using the Physical Activity Score for the Elderly (PASE). Locomotor function was measured as the time (in seconds) required to walk 20m. The twenty-meter walk test (20MWT) has been shown to be reliable when performed in a standardized fashion in a corridor free of obstructions and distractions. Functional mobility was measured as the time (in seconds) required for a person to stand from a seated position in a chair 5 times without using their arms.
SAS® was used for the analyses. The diagnosis of GTPS, as defined in the Physical Examination section, was treated as dichotomous for limb-specific analyses (present or absent) and trichotomous (neither, unilateral, or bilateral) for person-specific analyses. In addition to analyzing BMI as a continuous measure, a trichotomous categorical variable for BMI was defined using World Health Organization and National Institutes of Health definitions (25, 25 and <30, ≥30 kg/m²). Categorical variables were summarized using frequencies, proportions, and odds ratios (ORs); 95% confidence intervals (CIs) were calculated (sex, ITB pain, knee OA, knee pain, LBP). The Pearson chi-square test was used to compare proportions. Adjusting for age and sex as covariates and subject as a repeated-factor, limb-specific GTPS status was regressed on (1) ITB pain, (2) ipsilateral knee OA, and (3) contralateral knee OA using logistic regression with generalized estimating equations to adjust for correlation within subject. Adjusting for age and sex, person-specific GTPS status was regressed on (1) BMI as a continuous variable, (2) BMI category, and (3) LBP using logistic regression with generalized estimating equations to adjust for correlation within subject. Then, we assessed the association between limb-specific GTPS status and each of the possible risk factors while adjusting for age and sex, BMI no longer was associated with GTPS status with an OR of 2.19 (95% CI, 2.22–3.50). Similarly, functionally significant LBP that was significantly associated with GTPS, with an OR of 1.72 (95% CI, 1.34–2.19). Age and race were not found to be significantly associated with GTPS status (see table 1).

**Prevalence of GTPS**

A total of 5735 lower limbs from 2954 subjects were eligible for analysis of GTPS (fig 1). The subjects’ mean age ± SD was 62.4±8.1 years, and 60.1% were women. Of these subjects, 517 (17.6%) had GTPS. Specifically, 344 had unilateral and an additional 173 had bilateral GTPS, a prevalence of 11.7% and 5.9%, respectively (table 1). The prevalence of unilateral and bilateral GTPS was 15.0% and 8.5% in women and 6.6% and 1.9% in men. Comparing subjects with and without GTPS, the OR for GTPS in women compared with men was 3.32 (95% CI, 2.63–4.19). Age and race were not found to be significantly associated with GTPS status (see table 1).

**Biomechanic Correlates**

**ITB tenderness.** Logistic regression adjusting for age and sex revealed that ipsilateral ITB tenderness was significantly associated with GTPS, with an OR of 2.54 (95% CI, 2.03–3.17). After multivariate adjustment for all other variables, ITB tenderness continued to be significantly associated with GTPS with an OR of 1.72 (95% CI, 1.34–2.19). 

**Knee OA.** Controlling for age and sex, ipsilateral knee OA was significantly associated with GTPS, with an OR of 4.33 (95% CI, 3.43–5.48). Contralateral knee OA was also significantly associated with GTPS status with an OR of 2.19 (95% CI, 1.66–2.88). After multivariate adjustment for all other variables, ipsilateral (OR=3.47; 95% CI, 2.72–4.42) and contralateral (OR=1.74; 95% CI, 1.32–2.28) knee OA continued to be significantly associated with GTPS.

**Body mass index.** Logistic regression of GTPS on BMI, adjusted for age and sex, revealed that BMI was significantly associated with GTPS (P<.001). Compared with subjects with a BMI less than 25 kg/m², the OR for GTPS was 1.34 (95% CI, 1.09–1.66) in those with a BMI of 25 or greater and less than 30 kg/m² and 1.54 (95% CI, 1.15–2.07) for those with a BMI of 30 kg/m² or higher. However, after adjustment for ITB tenderness, ipsilateral and contralateral knee OA, LBP, as well as age and sex, BMI no longer was associated with GTPS status with an OR of 1.10 (95% CI, 0.80–1.52) comparing BMI equal to or greater than 30 kg/m² with a BMI less than 25 kg/m².

**Low back pain.** The presence of LBP most or all of the time was significantly associated with GTPS with an OR of 3.44 (95% CI, 2.76–4.28) after adjustment for age and sex. After adjustment for all other variables in the complete model, the association remained significant with an OR of 2.79 (95% CI, 2.22–3.50). Similarly, functionally significant LBP that

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>GTPS Case Subjects</th>
<th>Control Subjects</th>
<th>Intergroup Comparisons (P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age ± SD (y)</td>
<td>62.8±8.0</td>
<td>62.3±8.1</td>
<td>.198</td>
</tr>
<tr>
<td>Sex (% women)</td>
<td>80.7</td>
<td>55.7</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Race (%)</td>
<td>White</td>
<td>82.6</td>
<td>83.4</td>
</tr>
<tr>
<td></td>
<td>African American</td>
<td>15.3</td>
<td>15.3</td>
</tr>
<tr>
<td></td>
<td>All others</td>
<td>2.1</td>
<td>1.3</td>
</tr>
</tbody>
</table>

**Fig 1. Subject distribution.**

**Table 1: Demographic Distribution of GTPS**
limited activities over the past 30 days was associated with GTPS with an OR of 3.15 (95% CI, 2.54–3.92).

**Impairments**

Mean hip internal rotation ROM did not differ between hips with GTPS (36.7° from neutral; 95% CI, 36.1°–37.3°) and without GTPS (36.8° from neutral; 95% CI, 36.5°–37.1°) in those who underwent physical examination (P=.731). The presence of self-reported knee pain was associated with GTPS status with an age- and sex-adjusted OR of 2.75 (95% CI, 2.22–3.41).

**Person-Specific Function**

After adjustment for age and sex, GTPS did not significantly alter least mean square estimates of PASE score: no GTPS, 176.4 (95% CI, 173.1–179.6); unilateral GTPS, 169.0 (95% CI, 160.3–177.8); and bilateral GTPS, 171.9 (95% CI, 159.6–184.2). No significant difference was found among PASE in multiple comparisons comparing no GTPS with unilateral (P=.307) and bilateral GTPS (P=.790). Unilateral GTPS also did not alter 20MWT time. However, compared with subjects without GTPS, mean 20MWT time was 1.5±0.3 seconds greater in subjects with bilateral GTPS (P<.001). Similarly, the time to complete 5 sit-to-stand tasks was 0.9±0.2 seconds greater in subjects with unilateral and 1.5±0.3 seconds greater in subjects with bilateral GTPS (P<.001). Additionally, controlling for BMI did not change the overall associations detected.

**DISCUSSION**

This study was useful in both identifying the prevalence of GTPS in a non-clinic–based population as well as assessing the validity of common teachings regarding GTPS. The GTPS prevalence of 17.6% in this community-based sample of older adults at high risk of knee OA contrasts with the 20% to 35% reported for spine clinic patients presenting with LBP. We found the following were associated with GTPS: female sex, ITB tenderness, knee OA or knee pain, and LBP.

Reasons for the differences in prevalence may relate to the more strict definition of GTPS used in this study or the broader presence of self-reported knee pain was associated with GTPS status with an age- and sex-adjusted OR of 2.75 (95% CI, 2.22–3.41).

The increased odds for GTPS in women was consistent with prior reports, but we could not confirm a report associating GTPS with limited hip internal ROM. Although the mechanism for increased GTPS in women is unclear, this association could relate to anatomy (such as the flared pelvic rim in women altering the pull of the ITB), physiology (hormonal effects on bursal irritation or pain generators), or differences in activity between men and women.

The increased odds for GTPS with obesity and overweight status after adjustment for sex and age appeared to be explained by the combined effect of ITB tenderness, ipsilateral and contralateral knee OA, and LBP. The absence of an association between obesity and GTPS differed from prior reports, possibly because of the larger sample size and multivariate adjustment used in our study.

The associations with ITB pain, knee OA, and LBP also were consistent with orthopedic and physical medicine and rehabilitation teachings. Presumably, a tight ITB could explain tenderness on palpation over the lateral femoral condyle as well as potentially cause irritation over the greater trochanter. Additionally, knee or back pain may relate to GTPS through compensatory movements from 1 musculoskeletal problem causing symptoms at additional locations in the kinetic chain.

In addition to pain, people with GTPS appear to have a slowed gait and ability to rise from a chair. The impact on these functional activities may be significant to patients’ quality of life. However, this study did not detect an impact on participation in activities measured by the PASE instrument. The contrast between differences in function, but not in activity score, may relate to subjects with GTPS limiting their activities or to the reduced activity level in obese subjects.

A recent review of GTPS reiterated prior review articles in stating that “additional research is sorely required to shed light on . . . this commonly encountered syndrome.” This study was useful in investigating the validity of teachings about GTPS that previously had not been analytically assessed or proven. The study of GTPS in a community-based, rather than a clinic-based, sample was useful in clarifying the demographics of this condition.

**Study Limitations**

One limitation of this study was that the population studied was recruited for the MOST study because of characteristics that might predispose them to knee OA including obesity, knee injury, knee surgery, and knee pain. The presence of this recruitment bias may have influenced the prevalence of GTPS in our sample. However, this bias also may have enabled our findings to be more generalizable to community-dwelling adults who seek medical care. This may be more pertinent to clinical providers who serve patients with lower-body pain complaints.

Another limitation was that this cross-sectional study cannot establish causal relations. To elucidate whether ITB tightness, knee OA, obesity, or LBP lead to or are caused by GTPS, longitudinal studies would be necessary. The initial findings of this study may be useful not only in clinical care but also in generating hypotheses for mechanistic studies or therapeutic trials to benefit patients with GTPS.

**CONCLUSIONS**

The higher prevalence of GTPS in people who report hip pain in the absence of knee or generalized pain (24% of women, 9% of men) indicates that GTPS is common, and greater clinical awareness may identify patients for primary prevention and therapy. Slower functional performance in those with GTPS suggests that the study of targeted rehabilitation may be useful to minimize the impact of GTPS. A longitudinal study will be necessary to identify causal factors and outcomes of interventions.

**References**

5. Bird PA, Oakley SP, Shnier R, Kirkham BW. Prospective evaluation of magnetic resonance imaging and physical examination...

Suppliers
a. Ametek U.S. Gauge Div, Chatillon Brand Products, 8600 Somerset Dr, Largo, FL 33773.
b. Holtain Ltd, Crosswell, Crymych, Pembrokeshire SA41 3UF, UK.
c. SAS Institute Inc, 100 SAS Campus Dr, Cary, NC 27513.
The Lower-Limb Tasks Questionnaire: An Assessment of Validity, Reliability, Responsiveness, and Minimal Important Differences

Peter J. McNair, PhD, Harry Prapavessis, PhD, Jill Collier, MHSc, Sandra Bassett, PhD, Adam Bryant, PhD, Peter Larmer, MHSc


Objectives: To develop a questionnaire that focuses only on physical tasks related to lower-limb function and, within that questionnaire, to explore the psychometric properties of a series of questions that are related specifically to activities of daily living (ADLs) and a series of activities more often associated with recreation.

Design: Inception cohort.

Setting: Private practice.

Participants: Data were primarily from patients who had experienced a lower-limb injury that would typically involve rehabilitation up to 6 weeks.

Interventions: Not applicable.

Main Outcome Measures: Through 5 studies, the following psychometric qualities were evaluated: content and factor validity, construct and concurrent validity, test-retest reliability (intra-class correlation coefficient [ICC], typical error), responsiveness (effect size, standardized response mean, Guyatt’s responsiveness statistic), and the minimum important difference (distribution, anchor-based approaches).

Results: Factor analysis supported the theoretical perspective that ADLs and recreational activities can be treated as different domains within the construct of function. Internal consistency was high (Cronbach α: ADLs, .91; recreational activities, .95) and the 2 domains explained a moderate level of the response variance (61%). In the ADL domain, 7 tasks had greater than 80% of participants regarding them as having some importance. For recreational activities, 6 tasks had 79% or more of participants regarding them as having some importance. Both domains were moderately correlated to actual performance of tasks (r = .62, r = .72), and to other questionnaires used for lower-limb injuries (r range, .51-.86). The floor and ceiling effects of the domains followed an expected pattern that could be related to the loading forces experienced on the injured limb during activities. The reliability of the 2 domains was high (ICCs >.95), and the Bland-Altman plots showed that the distribution of error across the range of scores was random with low bias scores (<1.0 point). Typical error scores were 2 points for each domain. All measures of responsiveness were high (1.2–6.7). Measures of the minimal important difference varied (3–10 points) according to the methodologic approach used.

Conclusions: The questionnaire possesses good factor structure and composition, relates well with other measures of function, differentiates patients with regard to certain characteristics or processes known to occur after injury, shows high levels of reliability and responsiveness, and shows evidence of good minimal important difference scores. The findings support the use of the questionnaire in both clinical scenarios and in research.

Key Words: Leg injuries; Questionnaires; Rehabilitation; Reliability and validity.

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

The use of function as an outcome measure has been promoted by the World Health Organization (WHO). Function can be measured using self-report questionnaires that rely on the patient’s perception of their functional ability or, alternatively, it can be measured by directly observing the performance of actual tasks. Clinicians generally find the use of questionnaires to be more practical in the clinical environment where there is limited space for undertaking physical tasks. There are a multitude of self-report questionnaires available that are used to measure function and these can be categorized as generic, condition-specific, or patient-specific. Generic scales such as the Medical Outcomes Study 36-Item Short-Form Health Survey (SF-36) mix a wide range of concepts including emotional, mental, social, and physical function. These scales aim to measure the patient’s overall health status but may not provide a complete picture of a particular component such as function. Thus in some instances, the physical function subscale of these questionnaires may be less responsive to change than condition-specific questionnaires, making them less appropriate for use in a clinical setting where sensitivity to change might be of primary importance.

There are numerous condition-specific questionnaires that are used to measure different clinical disorders of the lower limb. This may present a problem for the clinician who is not a specialist, and treats many clinical conditions, because there is a requirement to store numerous questionnaires, have knowledge of their respective scoring systems, and know how to interpret those scores. To facilitate the use of outcomes related to function, a single questionnaire that covered a multitude of clinical problems in the lower limb would be most beneficial from a resource standpoint. Furthermore, with a single ques-
tionnaire, data can be more readily pooled by the clinician for both funding and accreditation organizations. Within the field of research, Deyo has also suggested that pooling of data would increase case numbers and allow additional statistical examination of the results from multiple small studies in which a treatment benefit has been observed yet statistical significance had not been achieved. It may also be important in comparing treatment responses of subgroups of patients across different studies. In achieving these objectives, the evidence base for rehabilitation would be improved.

From a theoretical perspective, it is important that the questions within a questionnaire remain focused on the construct of interest, in this case function. At this time, some questionnaires have included questions related to function and questions related to impairment (eg, pain, swelling), and their scoring involves summing of these quite different constructs. The rationale for such scoring is not clear. Other questionnaires have included general questions related to function (eg, ability to get into or out of a car). In doing so, they are mixing specific tasks that might fall within the tasks undertaken in the general question, and hence affect the validity and the responsiveness of the questionnaire. It is also apparent that some questionnaires have focused on function of daily living (ADLs) and activities related to work and sports, and these are scored by summing all questions. Thus, in many clinical situations, there will be questions that are suitable for some patients but certainly not for all. For instance, including questions for an elderly patient with a knee sprain that relate to their age, and perhaps unrelated comorbidities, which limits the amount of change that might be detected for the task being considered.

Also worthy of consideration is that at different stages of rehabilitation after an acute injury, it is likely that the difficulty associated with performing tasks related to recreation activity (eg, side stepping, jumping) is different compared with those tasks associated with ADLs, and furthermore the change in difficulty as rehabilitation progresses might also be different. In this respect, how many questions a questionnaire has related to ADLs and recreational activities may weight the questionnaire’s overall score and affect its ability to detect change in the patient’s overall function at particular times during the rehabilitation program. The separation of ADLs from recreational activities may solve some of these problems. However, at this time, no questionnaire examining musculoskeletal injuries has addressed the delineation of daily living and recreational domains.

Through 5 studies, this study describes the development of a Lower-Limb Tasks Questionnaire (LLTQ) that focuses only on specific tasks related to lower-limb function and explores the psychometric properties of a series of questions that are related specifically to ADLs and activities more often associated with recreation.

METHODS

The studies presented below were approved by the following Ethics Committees: Auckland University of Technology, Auckland Regional Ethics Committee, and the Central Queensland University of Technology. All subjects consented to participate.

Study 1: Content and Factor Validity

The LLTQ was initially developed from a review of literature that identified a number of questionnaires and scales used for lower-limb and spinal problems. Based on this review a list of 19 tasks used in daily activities, work, and sports was generated. This list of tasks was examined by a group of researchers and experienced clinicians who worked in the area of musculoskeletal medicine and rehabilitation together with a group of lay people. Their responses led to an additional 11 tasks being added to the initial set of questions.

The phrasing of the questions and the scale of measurement was then developed. WHO6 defines activity limitations as “difficulties” a person may have in executing activities. Because measures of difficulty have been examined closely by the Upper Extremity Collaborative Research Group who developed the Disability of the Arm, Shoulder and Hand questionnaire, it was thought that the same responses could be used in the LLTQ. Thus, for each task, subjects are asked to rate their difficulty to do the task in the past 24 hours. This choice of time period reflects a need to have patients recall information accurately, and it also appreciates the changes that occur in musculoskeletal injuries over short epochs. If subjects have not undertaken the task in the 24-hour period, then they are asked to make their best estimate on the response that would be most accurate. Subjects respond on a Likert scale using the following categories: no difficulty, mild difficulty, moderate difficulty, severe difficulty, and unable. These categories are scored from 4 to 0, respectively. To provide feedback concerning the phrasing and scale of measurement, 80 copies of the questionnaire were mailed to a sample of musculoskeletal treatment centers in the Auckland region. The practitioners in these centers (2–6 practitioners per center) were asked to provide a subjective assessment of the readability, understanding of the scoring, content, and ease of use of the questionnaire among their patients. Their responses led to some minor editing related to the format of the questionnaire to enhance the ease that it could be scored.

It was apparent from the list of activities in the LLTQ that some were related to tasks of daily living and others were more related to work and sporting activities. To establish whether these separate domains could be identified, a factor analysis was undertaken. This analysis also allowed an appreciation of how strongly each task was related to a particular domain. Using the above mentioned outpatient (nonhospital) treatment centers within the Auckland region, 200 copies of the LLTQ were mailed to clinicians who identified potential patients for inclusion. The criteria for inclusion were that participants were seeking treatment for a musculoskeletal disorder of the lower limb, were 18 years or older, and were able to ambulate normally prior to their injury. Patients were excluded if they had additional medical problems, which included cardiac and neurologic conditions, as well as amputation. The number of questionnaires that a center received was decided on by the owner or manager of the particular center, and over a 2- to 4-month period the questionnaires were completed.

To assess the LLTQ factor structure, a principal components analysis with oblique rotation was conducted on 159 responses to the LLTQ (see Table 1 for patient information). A preliminary assessment of psychometric adequacy was conducted (using the Bartlett test of sphericity: Kaiser-Meyer-Olkin sampling statistic) to determine the suitability of the LLTQ item correlation matrix for factor analysis. No restrictions were set limiting the number of factors in the analysis, so the items were free to load on any number of factors. Factor retention was determined based on eigenvalues greater than 1 (the Kaiser

Arch Phys Med Rehabil Vol 88, August 2007
Study 2: Construct and Concurrent Validity

Because there are no criterion standards for the measurement of function, a number of issues related to construct and concurrent validity were examined. Construct validity is apparent when an instrument behaves as would be expected in relation to another measurement and refers to the ability of a scale to differentiate subjects with regard to certain characteristics known to the disease process or injury in ways consistent with hypotheses. Concurrent validity is similar to construct validity but involves correlating an instrument with some criterion that is administered at about the same time. In the current study, the following 5 issues were assessed:

First, using data from 59 patients (mean age ± standard deviation [SD], 37 ± 15y) attending outpatient clinics (see table 1 for the subjects' information and inclusion and exclusion criteria as for study 1), the LLTQ correlated with other questionnaires used for lower-limb conditions (Lysholm Knee Rating Scale,5 Ankle-Hindfoot Scale,3 Cincinnati Knee Rating Scale,5 Patient-Specific Function Scale,11 Activity Scale.11 This scale consists of 6 physical activities: a 40-m walk, a 40-m run, figure-8 run over 12m, single-leg hop as far as possible ensuring the foot clears the ground, single-leg crossover hop over a 6-m course, and hopping up and down a flight of 14 steps. A dichotomous scoring system is used for each task (0, task was not attempted or not completed successfully; 1, task was successfully completed), and the scores of the individual tasks were summed to give a total motor activity score. Wilson et al11 reported that the Motor Activity Scale had high internal consistency (α = 0.9).

Fourth, using the inclusion and exclusion criteria above, floor and ceiling effects of the LLTQ were examined in 119 patients (mean age, 31 ± 14y) (see table 1 for patient injury information and inclusion and exclusion criteria as for study 1) presenting at outpatient clinics (N = 6). It was hypothesized that there would be minimal floor effects for the ADL domain as most patients reporting to outpatient clinics can undertake some ADLs. With respect to the recreational activities domain, it was hypothesized that there would be notable floor effects, as patients with moderate injuries would not be able to perform many of the high loading tasks required in this domain. The frequency of 0 scores and scores of 40 (maximum) were examined and the criterion for a notable effect was a frequency greater than 5%, a value thought to be reasonable.

Fifth, the importance of the tasks within the 2 domains of the LLTQ was assessed. The success of patient–treatment provider interactions is dependent on a shared understanding. Although 2 authors12,13 have commented that outcome measures should include elements that are considered by the patient to be of importance, there remains concern14 that the ability of outcome measures to truly reflect the patient’s views has been largely overlooked. In the current study, this aspect was explored by asking the patients to rate the importance of each of the tasks in the questionnaire (see questionnaire in appendix 1). The responses for this question were: not important, mildly important, moderately important, and very important. Using the study 1 inclusion and exclusion criteria, 123 patients (mean age, 31 ± 14y) (see table 1 for patient injury information) attending outpatient clinics with lower-limb conditions or injuries participated. In the analyses of these data, the percentages for the response “not important” were compared with the other categories which were grouped together and considered as “some importance.” An analysis of the frequency of these responses was then undertaken.
Study 3: Test-Retest Reliability

Seventy subjects (mean age, 33 ± 11 years) (see Table 1) completed the LLTQ on 2 occasions. The interval between occasions was 1 to 7 days. On the second occasion, the patients who had identified that they had not changed were asked to complete the questionnaire. The short duration between completion of responses was used to minimize the bias associated with true changes that might have occurred in their condition as a result of their recent visit to the clinician or natural healing processes. Descriptive statistics were calculated for the dependent variables. Data were plotted using Bland-Altman graphs enabling an appreciation of the distribution of error. Bland-Altman limits of agreement were also calculated. Typical error was calculated using the SD of the difference scores (day 2 minus day 1 responses). Intraclass correlation coefficients (ICCs) between the days were also calculated. In the latter, a 2-way mixed-model was used with the mode of assessment (days) as the fixed variable, and the subjects as the random variable.

Study 4: Responsiveness

The data used in this examination were from 174 patients (mean age, 34 ± 14 years) (see Table 1 for patient injury information and inclusion and exclusion criteria as for study 1) attending outpatient clinics (N = 9). They were treated for between 1 to 6 weeks. Additionally, a group of patients who had received an anterior cruciate ligament (ACL) reconstruction (n = 21) were assessed. Their inclusion allowed an appreciation of the responsiveness after significant lower-limb surgery and an examination of changes in the LLTQ over a greater period of rehabilitation: at 6 weeks postoperation and at 6 months postoperation. The following statistics were calculated for both groups: effect size, which was calculated as the mean at discharge from treatment subtracted from the mean at baseline divided by the SD at baseline; the standardized response mean, which uses the SD of the change in scores for the denominator instead of the baseline scores; and the responsiveness statistic, which is a measure of change relative to the variability in scores among subjects who are clinically stable, and requires the square root of twice the mean squared error of scores in the denominator.

Study 5: Minimal Important Difference

In respect to minimal important difference, 119 patients (mean age, 34 ± 13 years) (see Table 1) completed the LLTQ on 2 occasions. Initially, subjects completed the questionnaire at their first visit to the physical therapist. They then completed the questionnaire again 7 to 10 days after the initial visit. In addition to the LLTQ, subjects and the clinician completed a global impression of change. The global impression of change asked the following question: How would you rate your function since you last completed this form? The responses were: much better, better, same, worse, and much worse. Our a priori definitions of an important change were the categories: better or worse.

For each subject, the change score in the LLTQ in the ADL and recreation activities domains was calculated. To analyze these data, distribution-based and anchor-based methods were used. In respect to the former, SDs of the difference scores were divided by 3 to provide .33 SD estimates. This parameter reflects a change level between small and medium. The standard error (SE) of measurement was also calculated using the formula:

\[ \text{SE of measurement} = \sigma_e (1 - \alpha)^{1/2} \]

where \( \sigma_e \) is the SD of the LLTQ score and \( \alpha \) is the Cronbach \( \alpha \) coefficient, which is a measure of the internal consistency of the LLTQ domains (ADLs and recreational activity).

The anchor-based approach involved the global impression of change question. The LLTQ data from each subject were stratified into groups according to their responses to the global impression of change question, and descriptive statistics were calculated. A t test was used to determine whether there were significant differences between those subjects who regarded their condition as unchanged and those who reported that they were better. To further characterize the LLTQ changes and the impression of global change, receiver operating characteristic (ROC) curves were derived, and sensitivity and specificity values together with likelihood ratios measures were calculated. With respect to the latter measures, the area under the ROC curve relates to the overall ability of the questionnaire to correctly identify change versus no change in status. Specifically, the area indicates the probability that a subject who has stated that he/she had improved will have a higher score than one who states that he/she had not changed.

RESULTS

Study 1: Content and Factor Validity

The Bartlett test of sphericity relating to item interdependence was significant (\( \chi^2 = 2304.88, P < .001 \)) and the Kaiser-Meyer-Olkin sampling statistic of .92 was adequate, indicating that LLTQ item correlation matrix was suitable for the factor analysis. Twenty tasks were grouped into 2 coherent and interpretable factors related to ADLs and recreation activities (Table 2). These factors accounted for 61% of the variability in these items. Ten tasks did not load onto either factor. The Cronbach \( \alpha \) values for ADLs and recreational activities were .91 and .95, respectively, indicating that the constructs and domains of interest had good internal consistency. The 2 constructs correlated (r = .78), sharing approximately 61% common variance.

Study 2: Construct and Concurrent Validity

With respect to associations across questionnaires, moderate to high correlations (range, .51–.86) between the LLTQ ADL and recreation activities domains and the other questionnaires were observed (Table 3).

A comparison of the scores for the LLTQ and the PSFS showed that significantly higher scores were observed for the ADL domain of the LLTQ (effect size, 1.49), but there was no significant difference between the recreational activities domain and the PSFS (effect size, .07) (Fig 1).

In the comparison between physical performance and the LLTQ scores, moderate correlations between the Motor Activity Scale and the ADL domain (r = .62), and the recreational domain (r = .72) were observed. These correlations differed significantly from one another (r = 2.32, P < .05).

With respect to floor effects, the findings showed that 1.7% had a zero score in the ADL domain and 17% had a zero score in the recreational activities domain. With respect to ceiling effects, 9.8% had a score of 40 in the ADL domain, and no subjects had this score in the recreational activities domain.

The findings related to the importance of activities in the 2 domains of the LLTQ are presented in Table 2. It shows that the majority of participants regarded all tasks to have some importance. In the ADL domain, 7 tasks had greater than 80% of participants regarding them as having some importance. For recreational activities, 6 tasks had 79% or more of participants regarding them as having some importance.
Abbreviations: AHS, Ankle-Hindfoot Scale; CKQ, Cincinnati Knee Rating Scale; LYSH, Lysholm Knee Rating Scale; RA, recreation activities.

NOTE. All values are statistically significant at \( P < .05 \).

Abbreviation: NA, not applicable.

Study 3: Test-Retest Reliability

The results related to the Bland-Altman analyses are presented in figure 2A and B. They show that the distribution of the error was low and randomly distributed across the range of scores for ADLs and recreational activities. For the ADL, the bias and limits of agreement were \(-.63\) (range, \(-5.8\) to \(4.5\)) and for the recreational activities, \(-.52\) (range, \(-5.5\) to \(4.5\)). The typical error was 2 points for both the ADL and recreational activity domains. The ICCs and their lower confidence intervals (CIs) were \(.96\) (.93) and \(.98\) (.97) for the ADL and recreational activity domains, respectively.

Study 4: Responsiveness

For the ADL domain, for subjects attending outpatient clinics, the effect size, standardized response mean, and the responsiveness statistic were \(1.5\), \(1.4\), and \(5.3\), respectively. For the recreational activities domain scores, the same statistics were \(1.6\), \(1.3\), and \(5.8\), respectively. With regard to the ACL reconstruction patients, for the ADL domain the effect size, standardized response mean, and the responsiveness statistic were \(1.2\), \(1.4\), and \(3.0\), respectively. For the recreational activities domain scores, these same statistics were \(3.9\), \(2.6\), and \(6.7\), respectively.

Study 5: Minimal Important Difference

There were insufficient data to examine the category “worse”; therefore what follows relates to the category “better.” With respect to the distribution-based techniques, the \(.33\) SD scores for the ADL and recreational activity domains were 2.6 and 2.2 points, respectively. The SEs of measurement for these domains were 2.8 and 2.2, respectively.

With respect to the findings for the anchor-based approach, the mean change score \(\pm SD\) for subjects reporting no change in their condition was \(0.6 \pm 1.8\) (95% CI, 0.11 to 1.0) for ADL and \(0.4 \pm 2.0\) (95% CI, 0.1 to 0.95) for the recreational activities domain. For those subjects reporting that their condition was better, the mean change score was \(10.1 \pm 7.9\) (95% CI, 7.1 to 13.0) for ADL and \(7.7 \pm 6.8\) (95% CI, 5.1 to 10.2) for the recreational activities domain.

The area under the ROC curves \(\pm SE\) was \(91 \pm 4.9\) (95% CI, .82 to .99) for the ADL domain and \(88 \pm 4.4\) (95% CI, .79 to .96) for the recreational activities domain. For ADL change scores between 1 and 4 points, sensitivity ranged from 93% to 83%, and specificity from 63% to 98%. The respective likelihood ratios were above 10 (likelihood ratio, 11) for change scores greater than 2.5 with the maximum likelihood ratio (likelihood ratio, 54) observed at a change score greater than 3.5 points. For recreational activities, for score changes between 1 and 4 points, sensitivity ranged from 90% to 75%, and specificity from 66% to 92%. The respective likelihood ratios were above 10 (likelihood ratio, 13) for score changes greater than 3.5 with the maximum likelihood ratio (likelihood ratio, 30) observed for a score change greater than 6.0 points.

Overall, the results for minimal important difference were positive and showed it to be relatively low in most instances. However, they also highlight that there are differences in minimal important difference scores according to the analysis technique used.

Table 2: Factor Analysis—Oblique—Rotated Factor Loadings and the Percentage of Patients Who Regarded the Particular Task as Having at Least Some Importance in Their Life

<table>
<thead>
<tr>
<th>LLTQ Tasks</th>
<th>ADL Factor</th>
<th>Activities of Recreation Factor</th>
<th>Percent Rating Task Important</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walk for 10min</td>
<td>.82</td>
<td>NA</td>
<td>98</td>
</tr>
<tr>
<td>Walk up and down 10 steps</td>
<td>.79</td>
<td>NA</td>
<td>97</td>
</tr>
<tr>
<td>Stand for 10min</td>
<td>.59</td>
<td>NA</td>
<td>89</td>
</tr>
<tr>
<td>Stand for a typical work day</td>
<td>.75</td>
<td>NA</td>
<td>84</td>
</tr>
<tr>
<td>Get on and off a bus on a car</td>
<td>.85</td>
<td>NA</td>
<td>54</td>
</tr>
<tr>
<td>Get up from a lounge chair</td>
<td>.73</td>
<td>NA</td>
<td>84</td>
</tr>
<tr>
<td>Push or pull a heavy trolley</td>
<td>.69</td>
<td>NA</td>
<td>62</td>
</tr>
<tr>
<td>Get in and out of a car</td>
<td>.81</td>
<td>NA</td>
<td>96</td>
</tr>
<tr>
<td>Get out of bed in the morning</td>
<td>.50</td>
<td>NA</td>
<td>94</td>
</tr>
<tr>
<td>Walk across a slope</td>
<td>.86</td>
<td>NA</td>
<td>74</td>
</tr>
<tr>
<td>Jog for 10min</td>
<td>NA</td>
<td>.82</td>
<td>84</td>
</tr>
<tr>
<td>Pivot or twist quickly while walking</td>
<td>NA</td>
<td>.83</td>
<td>84</td>
</tr>
<tr>
<td>Jump for distance</td>
<td>NA</td>
<td>.85</td>
<td>57</td>
</tr>
<tr>
<td>Run fast/sprint</td>
<td>NA</td>
<td>.77</td>
<td>79</td>
</tr>
<tr>
<td>Stop and start moving quickly</td>
<td>NA</td>
<td>.76</td>
<td>84</td>
</tr>
<tr>
<td>Jump upward and land</td>
<td>NA</td>
<td>.88</td>
<td>68</td>
</tr>
<tr>
<td>Kick a ball hard</td>
<td>NA</td>
<td>.78</td>
<td>67</td>
</tr>
<tr>
<td>Pivot or twist quickly while running</td>
<td>NA</td>
<td>.84</td>
<td>79</td>
</tr>
<tr>
<td>Kneel on both knees for 5min</td>
<td>NA</td>
<td>.55</td>
<td>65</td>
</tr>
<tr>
<td>Squat to the ground/ floor</td>
<td>NA</td>
<td>.74</td>
<td>82</td>
</tr>
</tbody>
</table>

Table 3: The Correlation Coefficients Across Different Rating Scales

<table>
<thead>
<tr>
<th>Questionnaires</th>
<th>LLTQ</th>
<th>SF-36*</th>
<th>SF-36†</th>
<th>PSFS§</th>
<th>PSFS‡</th>
<th>CKQ</th>
<th>LYSH</th>
<th>AHS</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADL</td>
<td>.86</td>
<td>.86</td>
<td>.55</td>
<td>.59</td>
<td>.69</td>
<td>.76</td>
<td>.72</td>
<td></td>
</tr>
<tr>
<td>RA</td>
<td>.82</td>
<td>.84</td>
<td>.51</td>
<td>.66</td>
<td>.83</td>
<td>.79</td>
<td>.59</td>
<td></td>
</tr>
</tbody>
</table>

NOTE. All values are statistically significant at \( P < .05 \).

Abbreviations: AHS, Ankle-Hindfoot Scale; CKQ, Cincinnati Knee Rating Scale; LYSH, Lysholm Knee Rating Scale; RA, recreation activities.

*SF-36 for ankle and foot conditions.
†SF-36 for knee conditions.
§PSFS for ankle and foot conditions.
‡PSFS for knee conditions.
DISCUSSION

A key finding of the current study was that the factor analysis supported our theoretical perspective that ADLs and recreational activities can be treated as different domains within the construct of function. Being able to delineate these 2 domains allows a more complete picture to be gained of a patient’s function. According to the patient’s injury and demographics (eg, age), clinicians can use either domain or both domains to show change as a result of an intervention. Clinicians can also appreciate any differences in progress made in the different domains at various stages of the rehabilitation program. No previous research has examined function in detail from this perspective. The findings showed that the psychometric qualities of the 2 domains appear to be sound. Internal consistency was high (Cronbach α = .91, α = .95) and the factor analysis results produced 2 distinct interpretable domains that explained a moderate level of the response variance (61%). In addition, although the 2 domains correlated (r = .78), they only shared approximately 61% common variance. To put this another way, there was 39% unexplained variance between the 2 domains. These data, taken in concert, suggest that the LLTQ items are tapping into 2 conceptually different aspects of lower-limb function. Further evidence of differences in these domains was the finding that the recreational activities domain of the questionnaire correlated moderately with actual performance of tasks as measured by the Motor Activity Scale,\(^1\) which includes very similar jumping and running activities, and this correlation was notably higher than that for the ADL domain.

The LLTQ correlated moderately with other questionnaires used for lower-limb injuries providing evidence for its concurrent validity. High correlations were observed for both domains of the LLTQ and the physical function section of the SF-36. These were most likely due to the latter questionnaire combining questions related to moderate and vigorous activities as well as asking general and specific questions concerning tasks. With respect to the PSFS, when standardized to a percentage, the scores of the LLTQ ADL domain were significantly higher whereas those of the recreational activities domain were not significantly different. This finding highlights the differences between the ADL and recreational activities domains and suggests that patients prioritize their recreational tasks in respect to importance.

Although the inclusion of categories related to importance of each task was included in the questionnaire to assist in establishing its validity, these categories could be used for weighting of tasks. In its simplest form, this might involve multiplying the importance score by the difficulty score for each task and then summing the results across tasks. Theoretically, this might lead to increased levels of validity and responsiveness of the questionnaire. These ideas have not been examined at this time, but they are worthy of further investigation.

The floor and ceiling effects followed an expected pattern that could be related to the loading forces experienced on the injured limb during activities. Within the ADL domain where loading levels are relatively lower, few subjects (1.7%) had a floor effect, but a notable ceiling effect (9.8%) was observed. Most people with musculoskeletal injuries who are attending an outpatient clinic can undertake some activities of daily living without difficulty, and many are able to do these activities with no difficulty. There was a notable floor effect for recreational activities (17%) indicating that these activities which involve significantly more

---

**Fig 1.** Comparisons of LLTQ ADL and recreational activity (Rec) domains with the PSFS. *P<.05.

**Fig 2.** The Bland-Altman distribution showing the difference in scores across days with the average of scores across days for the (A) ADL and (B) recreational activity domains.
loading on body structures were of more difficulty. It would be unusual for a patient to report that they had no difficulty associated with activities that place significant load on the body, and hence no subjects had a maximum score in the recreational activities domain at the start of treatment. The reliability of the 2 domains as measured by ICCs was above .93, and hence well above the criterion value of 0.8 suggested for clinical measures.22 The Bland-Altman plots (see fig 2) provided evidence that the distribution of error across the range of scores was random, whereas the low bias scores (<1.0) indicated that there was minimal systematic error occurring across days. The limits of agreement for both domains were between 4 and 5 points and typical error was 2 points. These measures are conceptually different. Typical error provides a measure of the variation in test-retest scenarios whereas the limits of agreement provide the interval in which a patient’s difference score would fall 95% of the time. In the current study when a patient who is unchanged in status is retested, their score has a 1 in 20 chance of being 4 to 5 points above or below their baseline score. Hopkins16 has argued that this level of chance is too stringent, and suggests that half the limits of agreement would be more reasonable.

The responsiveness of a questionnaire plays an important role in decisions concerning the efficacy of treatment.20 Previous work has shown that general health status questionnaires are less responsive than disease- or joint-specific questionnaires. These findings usually reflect the broader set of questions used in general health scales, some of which may not actually be related to the condition or the intervention being tested, and hence they do not change notably over the course of rehabilitation. An aim related to the current study was to produce a series of questions that could be used for the lower limb across different disorders and anatomical locations and yet retain adequate levels of responsiveness. In the current study, all measures of responsiveness were above 1.2 and as high as 6.7. Based on Cohen’s suggestions23 that small and medium effect sizes are 0.2 and 0.5 SDs, respectively, our findings indicate that both domains of the LLTQ are very responsive. With respect to the responsiveness statistic, much higher values were observed compared with the effect size and the standardized response mean. This finding reflects the denominator used in the calculations, it being the variability of patients who are stable, thus Wright and Young1 refer to this statistic as providing a signal to noise ratio.

Guyatt et al20 commented that responsiveness reflects a treatment effect, and not necessarily a minimal important difference in the patients’ status. In the current study, 2 approaches were used to examine the minimal important difference. With respect to the distribution approach, based on Cohen’s work22 concerning the magnitude of effect sizes, Eton et al24 argued that .33 SD provides a reasonable approximation of a minimal important difference which in the current study equated to score changes of 2.6 and 2.2 for the ADL and recreational activities domains. It should be noted that distribution approaches that use the SD only in their calculations have been criticized because they are subject to bias as a result of subject heterogeneity that may occur within samples. Thus Wyrwich et al21 have recommended the use of SE of measurement, which is sample-independent because it uses both the SD and a measure of reliability in its calculation and is relatively stable across all but the extremes of population’s scores and thus has greater generalizability. A further advantage of the SE of measurement is that it is measured in the original units of the outcome measure and is therefore easier to comprehend. In the current study, the Cronbach α was high for both ADL and recreational activities domains (α=.91, α=.95, respectively) indicating the unidimensionality of the data within each domain, and this contributed to the relatively low SE measurement of 2.8 and 2.2 points. With respect to the anchor-based approach, the minimal important differences were higher (10 and 8 points for the ADL and recreational activities, respectively). There are a number of possible reasons for these higher values. First, we did not use the response “slightly better” as our standard for change, but followed the definition used by Farrar et al25 that a change of clinical importance was a rating of better or much improved. Farrar argues that the concept of being better is conceptually reasonable and clinically relevant. Nevertheless, it may be that the use of an additional response (slightly better) would have provided a lower estimate of the minimal important difference. It should also be noted that the reliability of the global question was not analyzed.

The results from the ROC curves and the subsequent calculations related to sensitivity and specificity provided further information concerning the minimal important difference for the LLTQ. In this context, the sensitivity refers to percentage of people who changed that the LLTQ correctly identified as changed, and specificity refers to the percentage of patients who did not change and were correctly shown not to have changed. It will differ according to the cutoff score used. If a high threshold change score is chosen, then the specificity is increased but the sensitivity is decreased. Likelihood ratios incorporate measures of sensitivity and specificity, providing the increase in odds favoring a condition being present.22 In the context of the current study, a condition being present would mean a shift in the level of function of the patient (ie, no change vs better). Using the criterion of 10, a value regarded as large shift in the probability of change being present,26 the cutoff score for ADL domain was 2.5 points, and the cutoff for the recreational activities was 4.5 points. These values are very similar to those of the distribution-based minimal important difference scores. Thus, given the evidence presented, it seems reasonable to calculate the mean of the anchor, distribution and likelihood ratio approaches. This suggests that for both domains, the minimal important difference is 4 points. This finding is particularly encouraging because it shows that the LLTQ domains are able to detect relatively small changes in function. This might be of particular benefit in future research where small effects are expected from an intervention, or where only a small sample of subjects are available for study. An important area for future research is how the minimal important difference and responsiveness of the LLTQ compare with those of other questionnaires.

CONCLUSIONS

This study describes the development of a lower-limb task questionnaire that has distinct domains related to ADLs and recreational activities. The questionnaire possesses good factor structure and composition, relates well with other measures of function, differentiates patients with regard to certain characteristics or processes known to occur after injury, shows high levels of reliability and responsiveness, and has evidence of low minimal important difference scores. Furthermore the inventory is succinct, easy to administer, and score. The current findings are encouraging and support the use of the questionnaire in both clinical scenarios and in research. As with any questionnaire, its validity and usefulness in the clinical scene will be established through further comparison with existing questionnaires and continued use over time.

Arch Phys Med Rehabil Vol 88, August 2007
APPENDIX 1: LOWER-LIMB TASKS QUESTIONNAIRE

ACTIVITIES OF DAILY LIVING SECTION

Patient: ___________________ Date: _____________

INSTRUCTIONS

Please rate your ability to do the following activities in the past 24 hours by circling the number below the appropriate response.

If you did not have the opportunity to perform an activity in the past 24 hours, please make your best estimate on which response would be the most accurate.

Please also rate how important each task is to you in your daily life according to the following scale:

1. = Not important
2. = Mildly important
3. = Moderately important
4. = Very important

Please answer all questions.

<table>
<thead>
<tr>
<th>TASK</th>
<th>NO DIFFICULTY</th>
<th>MILD DIFFICULTY</th>
<th>MODERATE DIFFICITY</th>
<th>SEVERE DIFFICITY</th>
<th>UNABLE</th>
<th>IMPORTANCE OF TASK</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Walk for 10 minutes</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1 2 3 4</td>
</tr>
<tr>
<td>2. Walk up or down 10 steps (1 flight)</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1 2 3 4</td>
</tr>
<tr>
<td>3. Stand for 10 minutes</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1 2 3 4</td>
</tr>
<tr>
<td>4. Stand for a typical work day</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1 2 3 4</td>
</tr>
<tr>
<td>5. Get on and off a bus</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1 2 3 4</td>
</tr>
<tr>
<td>6. Get up from a lounge chair</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1 2 3 4</td>
</tr>
<tr>
<td>7. Push or pull a heavy trolley</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1 2 3 4</td>
</tr>
<tr>
<td>8. Get in and out of a car</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1 2 3 4</td>
</tr>
<tr>
<td>9. Get out of bed in the morning</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1 2 3 4</td>
</tr>
<tr>
<td>10. Walk across a slope</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1 2 3 4</td>
</tr>
</tbody>
</table>

TOTAL (/40): ______

RECREATIONAL ACTIVITIES SECTION

Patient: ___________________ Date: _____________

INSTRUCTIONS

Please rate your ability to do the following activities in the past 24 hours by circling the number below the appropriate response.

If you did not have the opportunity to perform an activity in the past 24 hours, please make your best estimate on which response would be the most accurate.

Please also rate how important each task is to you in your daily life according to the following scale:

1. = Not important
2. = Mildly important
3. = Moderately important
4. = Very important

Please answer all questions.

<table>
<thead>
<tr>
<th>TASK</th>
<th>NO DIFFICULTY</th>
<th>MILD DIFFICULTY</th>
<th>MODERATE DIFFICITY</th>
<th>SEVERE DIFFICITY</th>
<th>UNABLE</th>
<th>IMPORTANCE OF TASK</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Jog for 10 minutes</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1 2 3 4</td>
</tr>
<tr>
<td>2. Pivot or twist quickly while walking</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1 2 3 4</td>
</tr>
<tr>
<td>3. Jump for distance</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1 2 3 4</td>
</tr>
<tr>
<td>4. Run fast/sprint</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1 2 3 4</td>
</tr>
<tr>
<td>5. Stop and start moving quickly</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1 2 3 4</td>
</tr>
<tr>
<td>6. Jump upwards and land</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1 2 3 4</td>
</tr>
<tr>
<td>7. Kick a ball hard</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1 2 3 4</td>
</tr>
<tr>
<td>8. Pivot or twist quickly while running</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1 2 3 4</td>
</tr>
<tr>
<td>9. Kneel on both knees for 5 minutes</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1 2 3 4</td>
</tr>
<tr>
<td>10. Squat to the ground/floor</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1 2 3 4</td>
</tr>
</tbody>
</table>

TOTAL (/40): ____
References


Quantitative and Clinical Description of Postural Instability in Women With Breast Cancer Treated With Taxane Chemotherapy

Meredith A. Wampler, PT, DPTSc, Kimberly S. Topp, PT, PhD, Christine Miaskowski, RN, PhD, Nancy N. Byl, PT, PhD, Hope S. Rugo, MD, Kate Hamel, PhD


Objective: To describe the postural control of women who received taxane chemotherapy for treatment of breast cancer using quantitative and clinically feasible measures.

Design: Prospective descriptive study.

Setting: University-based comprehensive cancer center.

Participants: Twenty women who completed taxane treatment for breast cancer and 20 healthy controls participated in this study.

Interventions: Not applicable.

Main Outcome Measures: Two quantitative measures of postural control were used, Sensory Organization Test (SOT) and center of pressure (COP) velocities. Two clinically feasible measures of postural control were used, the Fullerton Advanced Balance Scale (FABS) and Timed Up & Go (TUG) test.

Results: Compared with healthy controls, women with breast cancer had poorer postural control on all of the outcome measures. FABS and TUG scores correlated moderately with SOT and COP scores.

Conclusions: After taxane chemotherapy, women with breast cancer show significantly increased postural instability compared with matched controls. Clinically feasible measures of postural control correlated with quantitative tests. These results suggest that these clinical measures may be useful to screen patients to determine who may benefit from rehabilitation.

Key Words: Chemotherapy; Docetaxel; Equilibrium; Musculoskeletal system; Neoplasms; Paclitaxel; Peripheral neuropathies; Rehabilitation.

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

Breast cancer is the most common cancer diagnosis in women in the United States, with 178,480 women expected to be diagnosed with the disease in 2007. Chemotherapy is often used to treat this type of cancer. In 1994, the U.S. Food and Drug Administration (FDA) approved paclitaxel (Taxol), a drug derived from the Pacific Yew tree, for the treatment of metastatic breast cancer. In 1999, the FDA expanded the approval to include women with locally advanced disease (eg, positive lymph nodes, but no distant metastases). However, because there are limited natural resources to produce paclitaxel, the FDA approved docetaxel (Taxotere), a man-made drug with chemical properties similar to paclitaxel, for the treatment of locally advanced or metastatic breast cancer in 1998. Abraxane (paclitaxel protein-bound particles for injectable suspension) was approved in January 2005. These drugs comprise the taxane class of chemotherapy agents and are often used in combination with anthracyclines and cyclophosphamide (Cytoxan) to treat breast cancer. A growing body of evidence suggests these chemotherapy agents may improve survivorship and decrease the risk of local recurrence of breast cancer. Therefore, as taxane use becomes increasingly common, it is important for patients and health care providers to understand the impact that side effects of these drugs may have on physical function in women with breast cancer.

The side effects of taxanes include allergic reactions; neutropenia, anemia, and thrombocytopenia; alopecia; joint and/or muscle pain; mucositis; nausea; and peripheral neuropathy. Many of these effects are decreased with the use of medications such as erythropoietin, pegfilgastrim, and anti-emetics. Although many promising drugs are being tested, no effective peripheral nerve protectants are available at this time. Over half of the patients who receive paclitaxel, docetaxel, or Abraxane develop decreased deep tendon reflexes, increased vibration thresholds, or dysesthesias and paresthesias in a stocking and glove pattern. These patients may also show decreased sensory nerve action potential amplitudes on electrodiagnostic testing. A subset of these patients develop a painful peripheral neuropathy. Peripheral neuropathy can have a significant impact on physical function. Numerous studies have documented that diabetic peripheral neuropathy is associated with postural instability. There is also an increased risk of falls in this patient population. These patients show greater step width, decreased step length, and decreased gait speed in low light conditions on irregular surfaces. Differences in the temporospatial aspects of gait on challenging surfaces were also observed in adults.
with diabetic peripheral neuropathy who fell compared with those who did not fall.23

Although peripheral neuropathies are reported in women with breast cancer who are treated with taxane chemotherapy,10,15,24 to date, no studies have described postural stability. Therefore the primary purpose of this prospective study was to describe the postural stability of women who were treated with paclitaxel or docetaxel for breast cancer compared with matched healthy controls. We hypothesized that after the final infusion of taxane therapy, women with breast cancer would present with postural instability compared with matched healthy controls. In addition, the relationships between quantitative and clinically feasible measures of postural control were evaluated.

METHODS

Participants

For the study, we recruited 20 women treated with taxane chemotherapy for breast cancer and 20 healthy women between the ages of 30 to 60. These women were matched on age, height, and weight. All women were screened for pre-existing vestibular, visual, somatosensory, orthopedic, and neurologic disease before entering the study. Women in the breast cancer group were included only if there was no evidence of central nervous system metastases. All women were required to have a corrected low-contrast visual acuity better than 20/60 and a corrected high-contrast visual acuity better than 20/40. These entry criteria were chosen because postural control is related to visual acuity, orthopedic problems, neurologic problems, and vestibular problems.25-28 Women in the breast cancer group had received 4 cycles of doxorubicin (Adriamycin) and cyclophosphamide (Cytoxan) prior to beginning taxane therapy. All participants signed a written informed consent that was approved by the institutional review board and the protocol review committee at the university-based comprehensive cancer center before participating in the study.

We asked women in both the breast cancer and the healthy control groups to complete 1 testing session. Women in the breast cancer group were asked to return and repeat all tests within 1 week of initial testing to establish the intrarater reliability of quantitative and clinically feasible measures of postural control. Women in the breast cancer group were tested within 30 days of their final taxane infusion and healthy controls were tested at the time of enrollment. All measures, except for the Fullerton Advanced Balance Scale (FABS), were performed by 1 physical therapist who was not blinded to participants’ health status. Measurement bias was low for the quantitative measures of postural control and the Timed Up & Go (TUG) test, because none of these tools require a subjective evaluation of performance of a task. However, measurement bias was controlled for the FABS as described below.

All participants completed measures of height, weight, and visual acuity; a questionnaire regarding medical history and medication use; and quantitative and clinically feasible measures of postural control, as described below. In addition, all women completed several quantitative measures of peripheral neuropathy, including the total neuropathy score, the modified total neuropathy score, quantitative touch thresholds, quantitative vibration thresholds, and nerve conduction studies. As previously reported, the modified total neuropathy score (TNS) was found to be a clinically feasible measure to evaluate the severity of taxane-induced peripheral neuropathy,15 and therefore, only mean modified TNS scores of breast cancer and healthy control groups are reported to indicate the severity of peripheral neuropathy (table 1).

Measures of Participant Characteristics

Height was measured in centimeters using a stadiometer. Weight was measured in kilograms using a calibrated scale. Medical history and medication use, including chemotherapy, were self-reported by participants using a questionnaire. If women reported they were using a medication with known effects on the central nervous system (ie, gabapentin [Neurontin], zolpidem [Ambien], venlafaxine [Effexor], fluoxetine [Prozac], or lorazepam [Ativan]) this was noted. No women were taking opioid medications. The total number of women using centrally acting medications in each group was recorded for data analysis. Medical history and medication use was confirmed by chart review for women in the breast cancer group.

We measured visual acuity using 2 Bailey-Lovie log MAR charts,29 one with high contrast letters and another with low contrast letters, and is reported as a log MAR score. Log MAR is an acronym for log10 of the minimal angle of resolution and was developed as an improved measure of visual acuity, given the limitations of the Snellen chart. As the letters on the chart become smaller, the angle of resolution also becomes smaller. Therefore the log MAR score represents the minimal angle of resolution which the person can clearly see. Participants stood 6.1m from the chart while keeping both eyes open. If the participant normally wore glasses, they were permitted to wear glasses during the entire testing session. The participant would read each line of progressively smaller letters until they reached a line where they made an error. The Log MAR score was calculated by

\[
\text{log MAR score} = \frac{1}{\text{log}_{10} \text{angle of resolution}}.
\]

| Table 1: Differences in Demographic and Clinical Characteristics Between Women With Breast Cancer (BC) and Healthy Controls (HCs) |
|-------------------------------------------------|---------------------|---------------------|---------------------|
| Variable                                        | BC Group (n=20)     | HC Group (n=20)     | Statistic (P)       |
| Age (y)                                         | 50.35 ± 9.34        | 49.60 ± 9.08        | T score = -2.6 (.798) |
| Height (cm)                                     | 165.15 ± 5.57       | 163.77 ± 5.54       | T score = -7.8 (.439) |
| Weight (kg)                                     | 68.19 ± 9.39        | 68.70 ± 14.60       | T score = -.12 (.901) |
| Medications acting on the CNS (n [%])           | 9 (45)              | 5 (25)              | z score = 1.36 (.175) |
| Visual acuity (log MAR score)                   | -0.31 ± 0.145       | -.041 ± 0.161       | T score = 1.48 (.147) |
| High contrast                                   | .221 ± 0.146        | .123 ± 0.097        | T score = 2.51 (.017)* |
| Low contrast                                    | 6.9 ± 3.01          | 1.15 ± 1.27         | T score = 7.88 (<.001)* |

Abbreviations: CNS, central nervous system; log MAR, log10 of minimum angle of resolution; SD, standard deviation.

*Values are considered statistically significant different (P <.05).

Arch Phys Med Rehabil Vol 88, August 2007
recording the log MAR score of the last complete line read minus .02 multiplied by the number of correct responses on the next smaller line. A log MAR score of 0 corresponds to 20/20 vision on a Snellen chart. A lower log MAR score indicates better visual acuity.

**Instrumentation and Data Acquisition for Quantitative Measures of Postural Control**

We collected center of pressure (COP) data by using a Kistler forceplate to determine the stability of participants in 4 static positions: eyes open with head straight (EOHS), eyes open with head back 40° (EOHB), eyes closed with head straight (ECHS), and eyes closed with head back 40° (ECHB). Sensory feedback parameters for each position are summarized in Table 2. This protocol is comparable with a protocol used to test the stability of patients with diabetic peripheral neuropathy. Participants were asked to complete three 30-second trials of each condition. A tracing of their self-selected foot positions.

<table>
<thead>
<tr>
<th>Testing Position</th>
<th>Visual Feedback</th>
<th>Somatosensory Feedback</th>
<th>Vestibular Feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kistler forceplate test</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Eyes open/head straight</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Eyes open/head back</td>
<td>+</td>
<td>+</td>
<td>Altered</td>
</tr>
<tr>
<td>Eyes closed/head straight</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Eyes closed/head back</td>
<td>-</td>
<td>+</td>
<td>Altered</td>
</tr>
<tr>
<td>SOT</td>
<td>Condition 1</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Condition 2</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Condition 3</td>
<td>Altered</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Condition 4</td>
<td>+</td>
<td>Altered</td>
<td>+</td>
</tr>
<tr>
<td>Condition 5</td>
<td>-</td>
<td>Altered</td>
<td>+</td>
</tr>
<tr>
<td>Condition 6</td>
<td>Altered</td>
<td>Altered</td>
<td>+</td>
</tr>
</tbody>
</table>

Abbreviations: Visual feedback: +, eyes open; -, eyes closed; altered, visual surround moves; somatosensory feedback: +, platform steady; altered, platform rotates in the sagittal plane; vestibular feedback: +, head straight; altered, head tilted back 40°.

We calculated ICCs to establish intrarater reliability for all measures that act on the central nervous system. The test of 2 conditions becomes progressively more difficult, beginning with condition 1 in which the participant keeps eyes open on a stable platform and a nonmoving visual surround. The test then progresses by removing visual feedback (by closing eyes), altering visual feedback (by moving the visual surround), or altering somatosensory feedback (by allowing the platform to rotate in the sagittal plane). NeuroCom software was used to calculate an equilibrium score for each condition and a total equilibrium score for all conditions. The equilibrium score, which can range from 0 to 100, represents the amount of sway of the participant during the trial; a lower score indicates more sway and a zero indicates a fall. Reliability in the participants with breast cancer was calculated using the same methods used to calculate forceplate ICCs (total SOT: ICC = .86; condition 1: ICC = .82; condition 2: ICC = .92; condition 3: ICC = .67; condition 4: ICC = .75; condition 5: ICC = .81; condition 6: ICC = .65).

**Clinically Feasible Measures of Postural Control**

Many clinics do not have a forceplate or a NeuroCom system. There is literature to support that both of these quantitative measures correlate to clinical tests of balance in other populations. Therefore, clinically feasible tests of balance were used to evaluate the relationships between these measures and quantitative measures of postural control. Participants completed the FABS and the TUG test.

The FABS is a relatively new task-based measure of balance, that includes standing with feet together and eyes closed, reaching forward to retrieve a pencil held at shoulder height, turning 360° in a right then a left direction, stepping up and over a 15.2cm (6-in) bench, tandem walking, standing on 1 leg, standing on foam with eyes closed, 2-footed jumping for distance, walking with head turns, and responding to an unexpected trunk perturbation. The FABS was chosen because the tasks challenge the sensory systems (ie, visual, somatosensory, vestibular) used for postural control and therefore may be more sensitive to balance problems in patients with chemotherapy-induced peripheral neuropathy, a primarily sensory neuropathy. The quality of the performance of each task is scored using standardized ordinal scoring criteria, with total scores that can range from 0 to 40 points. Higher scores indicate a better performance. The FABS test was video-taped and scored by 2 reviewers who were blinded to the participants’ health status. The total FABS score had high reliability between session 1 and session 2 in this cohort (ICC = .92).

The TUG test is a timed test of a person’s ability to stand from a chair, walk 3m, turn, and return to a seated position. It was chosen because of its moderate correlation with SOT equilibrium scores (Pearson r = −.49). It had high reliability in the current cohort (ICC = .88).

**Data Analysis**

We calculated ICCs to establish intrarater reliability for all dependent variables and to establish interrater reliability for the FABS. A Kolmogorov-Smirnov test was used to ensure that the data met the criteria of normality to allow the use of parametric statistics. Although participants in the breast cancer and healthy control groups were individually matched by age, height, and weight, a more conservative unpaired t test that assumed unequal variances was used to compare means in the breast cancer and healthy control groups for each measure of postural control and each patient characteristic, except medications that act on the central nervous system. The test of 2
proportions was performed to test for differences in the number of centrally acting medications. Pearson correlations were calculated to test for associations between the quantitative and the clinically feasible measures of postural control. Cohen $d$ statistics were calculated to determine effect sizes between groups, while controlling for differences in standard deviations (SDs). A $P$ value of less than .05 was considered statistically significant. A conservative Bonferroni adjustment was applied for multiple comparisons of postural control measures (family $\alpha=.05$; 4 comparisons for forceplate analysis, $P<.0125$ to be considered significant; 7 comparisons for sensory organization analysis, $P<.007$ to be considered significant). All statistical calculations were performed using Minitab software except ICCs were calculated using the Department of Obstetrics and Gynecology at the Chinese University of Hong Kong webpage (http://department.obg.cuhk.edu.hk/researchsupport/statstesthome.asp).

RESULTS

All dependent variables and patient characteristics met the criteria of normality ($P>.05$) using the Kolmogorov-Smirnov test except for 4 of the SOT conditions and 5 of the forceplate conditions. However, those that did not meet the criteria for normality were very close to the cutoff ($P$ within .02). Because parametric statistic tests are considered robust and can withstand minor variations from normality, $^{35}$ we have reported our data using parametric techniques.

Participant Characteristics

No significant differences were found between the 2 groups in age, height, weight, number of centrally acting medications, and high contrast acuity (see table 1). However, women in the breast cancer group had significantly poorer low contrast acuity ($P=.017$) and a mild, yet significant, peripheral neuropathy ($P<.001$).

Quantitative Measures of Postural Control

As illustrated in figure 1, COP velocities in all 4 testing positions were significantly higher for women with breast cancer compared with healthy controls ($P<.0125$). As the conditions increased in difficulty, COP velocities increased for both groups. The highest velocities and largest group mean differences were found in positions in which vision was occluded or there was altered vestibular feedback (eyes open with head straight: Cohen $d=1.01$; eyes open with head back: Cohen $d=1.33$; eyes closed with head straight: Cohen $d=1.28$; eyes closed with head back: Cohen $d=1.29$). In addition, the equilibrium scores of the SOT were significantly lower in women with breast cancer for all conditions except condition 1 (fig 2). When SDs are accounted for, the largest group mean differences were observed for conditions 2, 3, 5, and 6, where vision is either occluded or altered, thereby increasing the patient’s reliance on somatosensory or vestibular input for postural stability (SOT1: Cohen $d=−.74$; SOT2: Cohen $d=−1.46$; SOT3: Cohen $d=−1.37$; SOT4: Cohen $d=−1.05$; SOT5: Cohen $d=−1.57$; SOT6: Cohen $d=−1.30$).

Clinically Feasible Measures of Postural Control

Interrater reliability was high for the FABS ($ICC=.98$). Therefore, rater 1 and rater 2 scores were averaged for further statistical analyses. Women with breast cancer had significantly lower FABS scores compared with healthy controls, indicating a poorer performance (table 3). In addition, a large significant increase in mean TUG scores (Cohen $d=−.77$) was found in women with breast cancer compared with healthy controls (see table 3). The FABS correlated moderately with each quantitative measure of postural control (table 4). The TUG scores correlated with COP velocities in the eyes open with head back, eyes closed with head straight, and eyes closed with head back positions (see table 4).

DISCUSSION

This study is the first to describe differences in various measures of postural control in women being treated with taxane chemotherapy for breast cancer compared with matched healthy controls. As we hypothesized, women who were treated with taxanes showed poorer static and dynamic postural control and poorer scores on clinically feasible measures of postural control. These findings suggest that in addition to the symptoms of peripheral neuropathy, these patients may experience problems with balance and gait that need further evaluation and treatment by a physical therapist.

Problems with postural control may be partially explained by somatosensory changes secondary to taxane chemotherapy. In this sample, although the severity of peripheral neuropathy was relatively mild; the COP velocities from forceplate measures and SOT scores for conditions 1, 2, and 3 were similar to patients with severe peripheral neuropathy secondary to diabetes. $^{20,36}$ Moreover, as previously reported, the modified TNS, a measure of the severity of peripheral neuropathy, was moderately correlated with the total SOT score ($r=.66$, $P=.002$). $^{15}$
have been associated with cyclophosphamide. Given these effects, specifically increased incidence of cataract formation, cornea. Nonetheless, chemotherapy-induced ocular changes may have a vestibular component. Most notably, 2 of the peripheral neuropathy observed in the women with breast cancer. These medications cannot be excluded as possible contributors to postural instability.

It is possible that changes in the visual system may occur after chemotherapy. Although visual changes are not documented as a primary side effect of the taxanes, doxorubicin, or cyclophosphamide, significant difference in low contrast vision was found in the women with breast cancer compared with the healthy controls. In addition, several reports have documented acute spontaneous visual symptoms, such as scotoma (an area of lost or depressed vision surrounded by an area of more normal vision), photopsia (an appearance of flashes due to retinal irritation), and blurred vision, associated with paclitaxel and cyclophosphamide treatment. Late effects, specifically increased incidence of cataract formation, have been associated with cyclophosphamide. Given these symptoms have been documented in patients with and without peripheral neuropathy, it is not clear whether these visual symptoms are associated with peripheral neurotoxicity to the optic nerve, vascular changes to the retina, or changes to the cornea. Nonetheless, chemotherapy-induced ocular changes may have contributed to the postural instability observed in the women with breast cancer.

Although no published reports of taxane induced vestibular toxicity were found, other chemotherapy agents that cause peripheral neurophyathy, such as platinum compounds, are associated with vestibular toxicity. Our data suggest that the postural instability observed in the women with breast cancer may have a vestibular component. Most notably, the largest differences in low mean equilibrium scores occurred during SOT conditions 5 and 6, which occur on an unstable surface with either absent or conflicting visual feedback, respectively. It has been suggested that these conditions reflect the ability of a patient to appropriately use vestibular information to maintain postural stability. In addition, mean scores for the women with breast cancer were 16 points lower on condition 5 and 8 points lower on condition 6 compared with previously reported mean scores for patients with peripheral neuropathy and insulin-dependent diabetes mellitus. However, the scores were not as low as those of patients with bilateral vestibular hypofunction (BVH) (n = 24) (breast cancer SOT5, 44.90 ± 19.70; BVH SOT5, 9.84 ± 17; breast cancer SOT6, 49.00 ± 20.30; BVH SOT6, 10.52 ± 20.56). Interestingly, when the SDs for both cohorts were considered, there appears to be a subset of women with breast cancer with scores that approach those with BVH.

Scores on the clinically feasible measures of postural control (ie, FABS, TUG) were significantly poorer in women with breast cancer. Given that the FABS test correlated moderately with all quantitative measures of postural control and the TUG correlated moderately with 3 of 4 forceplate measures of postural control, it may be that these measures could serve as useful measures of postural control for clinicians who do not have forceplate or NeuroCom systems in their clinics. Although significant differences in the FABS and TUG tests were found between women with breast cancer and healthy controls, the differences were relatively smaller than those observed for the quantitative tests of postural control.

**Study Limitations**

This study has several limitations. First, it was a prospective descriptive study that examined only 1 time point (ie, at the completion of taxane therapy), rather than an evaluation of changes in postural control over time that would indicate whether postural instability improves after cessation of taxane chemotherapy. A longitudinal design would have also allowed the investigators to determine changes in postural control after doxorubicin and cyclophosphamide treatment and further describe the changes after receiving taxane therapy. Perhaps there is an additive affect of chemotherapy agents on somatosensory, visual, vestibular, or other central nervous systems related to postural stability. In addition, pretreatment measures of postural control were not obtained for the women with breast cancer. This limitation was partially overcome by matching the women with breast cancer to healthy controls on potential confounding variables associated with postural control such as age, height, and weight. In addition, participants with known visual, vestibular, orthopedic, or neurologic dysfunction were excluded from this study. Other limitations include the relatively small sample size and multiple measures of postural stability. However, we attempted to control for these limitations by using conservative statistical methods such as unpaired t tests and Bonferroni adjustments. Finally, comprehensive measures of vision and vestibular function were not obtained on the study participants. However, vestibular and visual impairments are not considered primary side effects of

**Table 3: Differences in Scores on Clinically Feasible Measures of Postural Control in Women With Breast Cancer and Healthy Controls**

<table>
<thead>
<tr>
<th>Clinically Feasible Measure of Postural Control</th>
<th>BC Group (n=20) (mean ± SD)</th>
<th>HC Group (n=20) (mean ± SD)</th>
<th>T Score (P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FABS score (0-40 points)</td>
<td>33.90±3.46</td>
<td>36.48±2.13</td>
<td>−2.83 (.008)*</td>
</tr>
<tr>
<td>TUG (s)</td>
<td>6.69±.994</td>
<td>5.85±.86</td>
<td>2.43 (.020)*</td>
</tr>
</tbody>
</table>

*Values are considered statistically significant different (P<.05).

**Table 4: Correlations Between Quantitative Measures and Clinically Feasible Measures of Postural Control in Women With Breast Cancer**

<table>
<thead>
<tr>
<th>Quantitative Measures</th>
<th>FABS</th>
<th>TUG</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOT composite score</td>
<td>.581</td>
<td>.907*</td>
</tr>
<tr>
<td>COP velocity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eyes open/head straight</td>
<td>−.581</td>
<td>.907*</td>
</tr>
<tr>
<td>Eyes open/head back</td>
<td>−.541</td>
<td>.914*</td>
</tr>
<tr>
<td>Eyes closed/head straight</td>
<td>−.523</td>
<td>.918*</td>
</tr>
<tr>
<td>Eyes closed/head back</td>
<td>−.496</td>
<td>.926*</td>
</tr>
</tbody>
</table>

*Values are considered statistically significant different (P<.05).
doxorubicin, cyclophosphamide, or the taxane class of chemotherapies \(^{4,6,37,38}\) and therefore exhaustive measures of these impairments were not included. In future studies, it will be important to include more specific measures of visual and vestibular function. It will also be important to perform measures before and after each cycle of chemotherapy to determine whether there is an additive effect of chemotherapy agents on postural control versus significant visual and/or vestibular toxicity related to taxane chemotherapy alone.

**CONCLUSIONS**

This study is the first to identify significant changes in postural stability in women treated with taxane chemotherapy for breast cancer compared with matched healthy controls. Although peripheral neuropathy is monitored as part of standard care in these patients, postural stability is not traditionally measured. The FABS and TUG are sensitive to changes in postural stability and moderately correlate with quantitative measures which may make monitoring for postural instability more feasible in the clinical setting. More information about chemotherapy-induced effects on the sensory, visual, and vestibular systems will facilitate our understanding of the origin, prognosis, and management strategies to restore postural control in this patient population.

**Acknowledgments:** We thank Pooja Maniar, DPT, and Hope DeLevega, MSPT, for their assistance in scoring the FABS.

**References**


Suppliers
a. Multimedia Center, School of Optometry, University of California, Berkeley, CA 84720.
b. Type 9287A; Kistler Instruments Inc, John Glenn Dr, Amherst, NY 14228-2171.
c. Dynatronics Corp, 7030 Park Centre Dr, Salt Lake City, UT 84121.
d. The MathWorks Inc, 3 Apple Hill Dr, Natick, MA 01760-2098.
e. NeuroCom Smart Balance; NeuroCom International Inc, 9570 SE Lawnfield R, Clackamas, OR 97015.
Changes in Postural Control in Hemiplegic Patients After Stroke Performing a Dual Task

Laurent Bensoussan, MD, Jean-Michel Viton, MD, PhD, Marco Schieppati, MD, Hervé Collado, MD, Virginie Milhe de Bovis, MD, Serge Mesure, PhD, PT, Alain Delarque, MD


Objective: To determine the effects of an attentional task on hemiplegic patients’ postural control performances.

Design: Retrospective study.

Setting: Department of physical and rehabilitation medicine at a university hospital.

Participants: Twenty-three hemiplegic patients and 23 healthy age- and sex-matched control subjects.

Interventions: Not applicable.

Main Outcome Measures: Sway area and sway path of the center of pressure were measured during 30 seconds in standing subjects and patients under 3 conditions: eyes open (EO), EO while performing a simple arithmetic task (EO-AT), and eyes closed (EC).

Results: In the hemiplegic patients, the body sway area increased significantly with EC (P < .001) and during the EO-AT task (P < .017) in comparison with EO. Sway area with EO-AT remained, however, significantly smaller than with EC (P < .014). In the healthy subjects, the body sway did not differ significantly between the EO-AT and EO tasks (P < .42). The increase observed in the sway area and path in the hemiplegic population during the EO-AT task correlated significantly with age.

Conclusions: The postural performances of hemiplegic patients decreased during both the arithmetic task and the EC task. The cognitive task had no effect on healthy subjects’ postural performances. This study is the first to show the combined effects of age and dual task on the postural performances of hemiplegic subjects.

Key Words: Attention; Hemiplegia; Posture; Rehabilitation.

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

CHRONIC STROKE PATIENTS have to deal with problems resulting from postural control alterations. Many studies have shown that changes in the postural control process occur in hemiplegic patients. Hemiplegic patients show changes in the patterns of recruitment and delayed contraction of the paretic muscle during balance disturbances.1 Weight distribution on the lower limbs is asymmetric, in favor of the sound lower limb,2,5 and weight transfer from the sound to the hemiplegic limb is altered.3 Postural adjustments are also affected.2,10 Sway of the center of pressure (COP), as measured by a force platform, is greater in standing hemiplegic patients than in healthy subjects.2,4,5,11,12 Some studies have also shown the existence of greater visual dependence4,5 and altered vertical perception in hemiplegic patients.13-16

Postural control problems are more frequent in patients with right brain damage.13,17 Hemineglect may be responsible for balance disorders in hemiplegic patients after stroke.14 However, little is known about the interactions occurring between cognitive processes and the attentional demands involved in postural control in hemiplegic patients.

Postural control in healthy subjects is usually assumed to be an automatic task, but many studies using a dual-task paradigm have shown that postural control requires attention.18 To define attention, we quote the most frequently used definition, which was given by William James in 1890: “It is the taking possession by the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought. Focalization, concentration, of consciousness are of its essence. It implies withdrawal from some things in order to deal effectively with others...”19(p22) Attention is the first step in learning. Information processing could be subdivided into 2 processes: an automatic process and a controlled or conscious process. The controlled process may be perturbed if another task has to be performed simultaneously or if the task is too lengthy. Attention can be selective, focused or divided, phasic or sustained. Each type of attention can be assessed using different tests, such as daily life questionnaires or arithmetic tests or by measuring reaction time.

The aim of dual-task procedures is to assess the information processing capacity of subjects performing 2 tasks at the same time. The tasks can be said to be independent if each task is performed correctly, which shows that the attentional requirements of the 1 task are not necessary to perform the other task. However, if the performance level obtained in 1 of the 2 tasks decreases in comparison with that obtained when this task is performed alone, this means that the attentional requirements of the 2 tasks exceed the subject’s attentional capacity.20 Dual tasks have been used with many groups of subjects, such as lower-limb amputees21 and patients with traumatic brain injury22 performing a cognitive task (primary task) at the same time as a postural task (secondary task). If subjects’ attentional capacity is exceeded, their performance on 1 of the 2 tasks, or both, will deteriorate.

With the dual-task paradigm, some studies have shown that the attentional capacity of hemiplegic patients is lower than
that of healthy subjects. Other studies have shown that the attentional demands increase depending on the difficulty of the postural task and in fallers. In these studies, changes in the cognitive performances have been assessed during a postural task, whereas changes in postural performances were not assessed during the cognitive task. Another study has assessed the effects of rehabilitation on hemiplegic subjects’ balance capacity. This longitudinal study on patients before and after rehabilitation showed that the postural control processes changed in postacute stroke patients performing dual tasks with their eyes closed (EC) as compared with the eyes open (EO) condition, in terms of weight bearing and COP velocity. The hemiplegic patients were not compared, however, with healthy sex- and age-matched subjects. No studies have therefore been performed so far on postural control in chronic hemiplegic patients performing a cognitive task in comparison with healthy subjects.

We have also addressed the influence of age on postural control during the performance of a dual task by hemiplegic patients. Several studies on healthy older subjects have shown the occurrence of changes in the postural control of body sway under dual-task conditions and during gait. To our knowledge, however, the effects of age on dual-task performance have never been studied in hemiplegic subjects. The aim of the present study was therefore to determine whether any changes in the postural control processes occur in hemiplegic patients of various ages with chronic stroke performing a cognitive task, and what differences they may show in comparison with healthy subjects. The dual task used here was designed to establish the importance of attention in postural control, to obtain further information about the postural control processes, and to develop new tools for assessing postural control in hemiplegic patients. Studies on dual-task paradigms are of considerable clinical importance. Hemiplegic subjects frequently perform dual tasks in everyday life, such as standing and talking or standing and washing or standing and dressing. It is therefore necessary to determine what effects the performance of dual tasks may have on postural control. If the performance of a dual task can sensitize the dynamometric recordings, then this paradigm may also constitute a useful simple test for assessing the postural abilities of hemiplegic subjects in clinical practice.

METHODS

We used an AMTI forceplate measuring 0.6×0.4m to assess the displacement of the COP in hemiplegic patients and healthy subjects. The sampling rate was 100Hz. The AMTI-netforce software program was used for the data recording and the Bioanalysis software program was used for the data processing. The parameters measured were the sway area and the sway path. A GAITRite system was used to assess subjects’ walking speed, in order to characterize the 2 populations of hemiplegic and healthy subjects.

Participants

This retrospective study was carried out on hemiplegic patients undergoing postural assessment at our department who had a stroke at least 5 months previously. Assessments of this kind are carried out routinely on hemiplegic patients at our department. The inclusion criteria were: poststroke hemiplegia and the ability to maintain a standing position without any help for 30 seconds. The exclusion criteria were: vestibular disease, cerebellar disease, visual impairments not corrected by glasses, cognitive disorders making it impossible to perform the dual task, and aphasia.

We included 23 hemiplegic patients in this study. There were 13 men and 10 women (mean age ± standard deviation, 50.3±14.9y; range, 22–77y), including 12 left hemiplegic patients and 11 right hemiplegic patients. Four of these patients had suffered from hemorrhagic strokes and 19 from ischemic strokes. The mean time elapsing since stroke was 37±28 months (range, 5–110mo).

Hemiplegic patients were compared with 23 healthy age- and sex-matched control subjects (10 men, 13 women) with a mean age of 44.4±14.4 years (range, 24–78y).

Assessment Procedure

We performed a clinical assessment on each subject to quantify the lesion, neurologic and joint impairments, hemineglect, balance, and gait impairments. The Fugl-Meyer Assessment modified by Lindmark and Hamrin was used to assess balance control. Sensory impairments were clinically assessed to detect any epicritic and proprioceptive sensory deficits. The pick and touch test and graphesthesia test were used for epicritic sensory assessment, and the arthrokinesia test and pallesthetic test for proprioception assessment.

Quantitative assessments of balance were carried out under the following conditions. The subjects were asked to adopt the most stable standing position, barefoot, on a piece of hardback paper fixed to the forceplate. The foot position chosen by the subjects was drawn on the hardback paper to ensure that the same position would be adopted by the subject under the various test conditions. The recording session lasted for 30 seconds. The first test was run in the EO condition. After a break of 15 seconds, the second test was run in the EC condition. After a further break of 15 seconds, the third test was run in the eyes open with an arithmetic task (EO-AT) condition. The arithmetic task was a 1-digit count-down from 50 (50, 49, 48, . . .). During the EO tasks, the subjects were asked to fix a target placed in front of them at eye level. Before each recording began, subjects were asked not to talk (except in the dual task), to keep their arms alongside their trunk, and not to move. In the dual task, subjects were asked to perform the task while counting loudly at a normal talking speed to check the execution of the dual task.

To assess the body weight distribution, we used a second AMTI forceplate. The weight distribution between the lower limbs was measured in the hemiplegic patients in the standing position with their eyes open, with 1 leg on each AMTI forceplate. The 2 forceplates were used only to determine the body weight distribution, and not during the performance of the dual or eyes closed tasks.

Data Analysis

The displacement (sway path length [in centimeters]) and the sway area of the foot COP (the 95% confidence ellipse drawn on the distribution of the data points [in cm²]) were analyzed during the 30-second recording epoch. These data were analyzed under the 3 conditions EO, EC, and EO-AT. The right and the left hemiplegic patients were also compared separately under the same conditions.

Statistical Analysis

We used the Student t test to check the differences between the 2 populations. Analysis of variance (ANOVA) was performed on all the other comparisons. EO sway path was compared with the EC path and EO-AT, and the EC sway path was compared with the EO-AT path in both hemiplegic and healthy.
subjects. Similar comparisons were carried out on the sway areas. The results obtained on hemiplegic patients in the 3 conditions were also compared to those obtained on healthy subjects under the same conditions. Linear regression was performed to compare the effects of age on sway area and path between healthy subjects and hemiplegic patients. The data subjected to the ANOVA were normally distributed (as confirmed using the Kolmogorov-Smirnov test).

RESULTS

Population

In the hemiplegic patients’ group, there were 12 subjects with superficial sensory impairments and 10 subjects with proprioceptive sensory impairments on the hemiplegic side (test with arthrokinesia), and 4 subjects showed hemineglect in the Bells test.30 This population included 1 patient with type 2 diabetes without any neuropathy at the clinical examination. This patient was being treated with oral medication for the diabetes. No subjects had any clinical signs of neuropathy at the clinical examination performed on the sound limb. There were 3 fallers scoring less than 1 fall a week. The mean value of the hemiplegic patients’ self-selected walking speed was 4.22±.23m/s, as compared with that of the healthy subjects, which was 1.16±0.14m/s. The weight distribution in hemiplegic patients was asymmetric, and significantly in favor of the sound leg in comparison with the hemiplegic leg (P<.001).

Sway Area and Path in Hemiplegic Patients

The sway area was greater in the EC condition than in the EO condition in hemiplegic patients (P<.001). The sway area was also greater in the EO-AT condition than in the EO condition (P<.017). The sway area differed significantly between conditions EO-AT and EC in hemiplegic patients (P<.014) (fig 1).

The sway path was longer in the EC condition than in the EO condition in hemiplegic patients (P<.001). The sway path length differed significantly between conditions EO-AT and EO (P=.05). The sway path was longer in the EC condition than in the EO-AT condition (P<.01) (see fig 1).

Sway Area and Path in Healthy Subjects

The sway area was larger in the EC condition than in the EO condition (P<.036) in healthy subjects. The sway area did not differ significantly in these subjects between conditions EO-AT and EO (P<.42). The sway area was larger in condition EC than in condition EO-AT in healthy subjects (P<.01) (see fig 1).

The sway path did not differ significantly between conditions EC and EO (P<.12) in healthy subjects. These subjects’ sway paths did not differ significantly between conditions EO-AT and EO (P<.58), nor did their sway paths differ significantly between conditions EC and EO-AT (P<.15) (see fig 1).

Comparisons Between Hemiplegic Patients Versus Healthy Subjects

Sway area in the EO condition did not differ significantly between hemiplegic patients and healthy subjects (P=.99). Sway area in the EC condition differed significantly between hemiplegic patients and healthy subjects (P=.016). A greater sway area was recorded in the EO-AT condition in hemiplegic patients than in healthy subjects (P=.004) (see fig 1).

Sway path was longer in the EO condition in hemiplegic patients than in healthy subjects (P=.03). Sway path was longer in the EO-AT condition in hemiplegic patients than in healthy subjects (P<.001). Sway path was also longer in the EC condition in hemiplegic patients than in healthy subjects (P<.001) (see fig 1).

Effects of Age on Dual-Task Performances

In hemiplegic patients, a significant increase in the sway area (P=.05) was found to occur with age in the dual-task condition. The slope of the linear regression was significantly steeper in hemiplegic patients than in healthy subjects (r=.1208, P<.001) (fig 2). In hemiplegic patients, a significant increase in sway path length (P<.01) was found to occur under the dual-task conditions with age. The slope of the linear regression was significantly steeper with hemiplegic patients than with healthy subjects (r=.724, P<.001) (see fig 2). The sway path length and sway area of the hemineglect patients were also compared with those of the other hemiplegic subjects to check.
whether hemineglect was more prevalent with age, because this might be a factor contributing to the differences between the performances of older and younger patients. The 2 groups did not differ significantly in terms of age ($P = .727$), sway path ($P = .681$), or sway area ($P = .387$).

**Comparisons Between Left Versus Right Hemiplegic Patients**

Body sway area did not differ significantly between left and right hemiplegic patients ($P = .63$). Body sway path length did not differ significantly between left and right hemiplegic patients ($P = .28$) (fig 3).

**Weight Distribution in Hemiplegic Patients**

The weight distribution data were analyzed (fig 4) and compared between the left and right hemiplegic groups. In the left hemiplegic group, the asymmetry of the weight distribution was significantly greater than in the right hemiplegic group. In the left hemiplegic group, the weight bearing on the sound lower limb was greater than on the contralateral side ($P < .001$). In the right hemiplegic group, there was no significant difference in the weight distribution between the limbs ($P = .067$).

**DISCUSSION**

The results obtained in the present study show that changes in the postural control processes occur in hemiplegic patients with chronic stroke of various ages performing a cognitive task. Changes in postural control were found to occur in the patients performing attention-demanding cognitive tasks: the body sway area and path length, as recorded by a dynamometric platform, increased when hemiplegic patients were performing a dual task. This was not the case in healthy subjects. In addition, the effects of age on the dual-task performance were established: under dual-task conditions, body sway was greater in the older than in the younger hemiplegic patients.

The effects of dual task on postural control have been previously documented in patients with other pathologies as well as in young and older healthy subjects but not in chronic hemiplegic patients as compared with control subjects. Our results are in agreement with those of de Haart et al., who reported that a change in postural control occurred under dual-task conditions as compared with eyes open conditions. However, the aim of that study was to assess recovery of balance in postacute stroke patients and no comparisons were made with control sex- and age-matched subjects. To our knowledge, the present study is the first to assess postural control using posturographic methods in chronic stroke patients under dual-task conditions.

In the present study, the hemiplegic patients showed an increase in body sway area and path length in the EC condition as well as in the EO-AT condition. Nardone et al. have described this effect under EC versus EO conditions in hemic-
plegic patients in comparison with healthy subjects. These authors noted that the sway area was larger under EC than EO conditions in the case of the hemiplegic patients. This suggests that attention is as necessary as vision for postural control in hemiplegic patients.

The second important finding to emerge from this study concerns the effects of age on the dual-task performances of hemiplegic patients. The stance stability of older hemiplegic patients decreased in the dual-task condition in comparison with younger hemiplegic subjects and healthy subjects. Moreover, increasing age was found to be closely correlated with decreasing postural performances in hemiplegic subjects when they performed the arithmetic task. This was in keeping with the results of studies performed on healthy older subjects described by Woollacott and Shumway-Cook18 in their review.

However, no difference was found to exist between older and young healthy subjects performing a dual task. These results may be explained by the fact that the arithmetic task was a simple one. Indeed, Woollacott and Shumway-Cook18 reported that the effects of cognitive task performance on postural control in older subjects depend on many factors, including the complexity of the cognitive task, the difficulty of the postural task, and the age and balance abilities of the subjects. There seem, therefore, to be additional effects of age and hemiplegia making dual tasks more difficult to perform.

The dual task seems to be a useful paradigm for assessing postural control in hemiplegic patients. Using a simple cognitive task such as the present arithmetic task, we established that a change occurs in hemiplegic patients’ postural control as the result of the attentional demands imposed. This cognitive task is one that can be easily performed by most hemiplegic patients. The performance of tasks of this kind had no effect on the healthy subjects.

There are several possible explanations for the change in postural control occurring in hemiplegic patients during the performance of a cognitive task. In the first place, their attentional capacity may have decreased.5,33,32 Marshall et al.31 have reported that hemiplegic patients performing a visual task showed an impaired attentional capacity, as shown by the lengthening of the dual-task performance times. Stapleton et al.31 have reported that the sustained and selective auditory attention and the selective visual attention are affected in hemiplegic patients and that some patients also showed visual inattention. Many studies, such as that by LaBerge,20 have mapped the brain location of attentional processing. The brain regions responsible for processing attention are the prefrontal, posterior cortical, and thalamic regions. Van der Werf et al.32 have reported attention disorders in stroke patients with complex attention deficits after damage to the intralaminar nuclei.

Another explanation could be that the attentional demands involved in performing cognitive or postural tasks increase in hemiplegic patients. For example, the equinus varus foot of a hemiplegic subject was responsible for difficulties in maintaining the standing position, possibly due to an increase in the attentional demands involved in maintaining equilibrium. This hypothesis is more difficult to check. The third possible explanation is that under dual-task conditions, a decrease in attentional capacity may be associated with an increase in the attentional demands involved in performing the postural control task and the concomitant cognitive task.

In this study, we did not observe the existence of any difference in body sway between left and right hemiplegic patients. Rode et al.17 have reported a larger sway area and a greater displacement of the COP under the sound lower limb in left hemiplegic compared with right hemiplegic patients. Lesions affecting the right brain hemisphere may result in more severe postural control disorders in hemiplegic patients, and in more pronounced spatial data processing impairments. However, in that study, 11 subjects previously showed hemineglect that had disappeared at the moment of the assessment, but those patients may still have had persistent postural representation deficits. In our population, only 4 of the 12 left hemiplegic patients had hemineglect. Another possible explanation for the absence of any differences in body sway between left and right hemiplegic patients might be found in the methods used in the present study. Our hemiplegic patients were free to choose their most stable foot position. This probably reduced the importance of differences in the changes in postural control between the right and left hemiplegic patients’ groups. Another explanation for the lack of difference between left and right hemiplegic subjects is that the time elapsing between the acute stroke and the posturographic measurement was as long as 28 months on average, that is, quite a long time had elapsed since the acute event. The differences between these 2 groups probably decrease with time after stroke. However, the body weight distribution between the lower limbs was found to be more asymmetric in left than right hemiplegic patients; greater relative weight was placed on the sound lower limb in left than right hemiplegic subjects. This asymmetry of the body weight distribution has been observed in previous studies on hemiplegic subjects.13,17 In our study, the hemiplegic patients with right brain damage showed greater changes in their postural control (asymmetry of body weight distribution) than the hemiplegic patients with left brain damage. This finding is in agreement with the data available in the literature.

The present hemiplegic and healthy subjects were free to place their feet in the most stable position. The position of their feet was not imposed as in other studies in the literature, because many patients could not even stay in the standing position with their feet together. The freedom they were allowed to choose the position of their feet might explain why the same body sway area was recorded in hemiplegic patients and healthy subjects in the EO condition. Admittedly, the subjects were asked to count aloud during the dual-task condition. Some studies have mentioned changes in the body sway occurring when subjects are speaking.3,5 The respiratory and articulatory movements involved in speaking might account for these changes in the body sway. In our study we had to make sure that the arithmetic task was being performed, which is why we...
asked the patients and subjects to count aloud. However, performing the arithmetic task aloud did not affect the postural control parameters in our healthy subjects. It can therefore be concluded that this factor did not bias our study.

Dual tasks provide a useful means of testing postural control in hemiplegic patients in clinical practice, because this situation often occurs in everyday life, when a subject is standing and talking or standing and washing or standing and dressing, for instance. In addition, dual tasks provide a simple test that increases the sensitivity of the postural tests commonly used in clinical practice. Studies on dual-task paradigms are therefore of considerable clinical importance. It would be interesting to conduct further studies to assess the relationships between the changes in postural control occurring during the performance of dual tasks and the risk of falling to which hemiplegic patients are exposed. This useful paradigm also provides a simple means of assessing whether hemiplegic subjects showing good postural control in simple tasks condition may use different postural control strategies under dual-task conditions, showing the existence of latent postural control abnormalities. In fact, this study has shown that the postural control of hemiplegic patients is superimposable on that of control subjects under eyes open conditions, whereas dual tasks involve changes in hemiplegic patients’ postural control patterns. Dual task conditions therefore increase the sensitivity of the postural tests available for use on hemiplegic subjects. This paradigm could also be used to assess the effects of therapy on hemiplegic subjects. It could be used for example to show the effects of botulinum toxin injection (intended to reduce spasticity) on postural control. The dual task can be used here to assess the attentional demands involved in postural control: if the patient’s stability improves after botulinum injection, this means that the attentional demand has decreased.

**CONCLUSIONS**

The results of this study show that hemiplegic patients’ postural control performances deteriorated during the performance of a dual task. The effects of age on hemiplegic patients’ dual-task performances were also reflected in the increase in sway area and sway path length recorded in the cases of the older patients performing the dual task. The performance of the simple arithmetic task had no effect on healthy subjects, whereas the attentional capacity of hemiplegic patients did not suffice to perform the dual task, and the quality of their stance decreased accordingly. Dual tasks therefore seem to be a useful means of assessing postural control in hemiplegic patients.

The dual-task performances were also reflected in the increases in postural control. The dual task can be used here to assess the attentional demands involved in postural control: if the patient’s stability improves after botulinum injection, this means that the attentional demand has decreased.

**References**

32. Van der Werf YD, Scheltens P, Lindeboom J, Witter MP, Uylings HB, Jolles J. Deficits of memory, executive functioning and attention following infarction in the thalamus; a study of 22 cases with localised lesions. Neuropsychologia 2003;41:1330-44.

Suppliers

a. Advanced Mechanical Technology Inc, 176 Waltham St, Watertown, MA 02472.
b. CIR Systems Inc, 60 Garlor Dr, Havertown, PA 19083.
Predictability of Simple Clinical Tests to Identify Shoulder Pain After Stroke

Bala S. Rajaratnam, MAppSc (PT), N. Venketasubramanian, FRCP, Prem V. Kumar, FRCS, James C. Goh, PhD, Y.-H. Chan, PhD


Objective: To identify simple diagnostic musculoskeletal tests that can be performed early after stroke to predict patients’ likelihood of reporting early signs of hemiplegic shoulder pain.

Design: Case control.

Setting: Multicenter acute care hospitals.

Participants: A total of 152 adults after a first episode of stroke, of whom 135 met the inclusion criteria. Thirty patients were assigned to the experimental group because they reported moderate intensity of hemiplegic shoulder pain at rest. The remaining 125 patients made up the control group.

Interventions: Not applicable.

Main Outcome Measures: Therapists measured the performance of combined upper-limb movement including the hand–behind-neck (HBN) maneuver, passive pain-free ranges of shoulder motion, 3 musculoskeletal tests, and the strength of deltoid muscles during each patient’s hospital stay. The numeric rating scale (NRS) identified those who reported moderate or greater intensities of hemiplegic shoulder pain during rest and during assessment.

Results: In our study, 22.2% (95% confidence interval, 15.5–30.2) of the patients reported hemiplegic shoulder pain, on average 1 week after the onset of stroke. Positive Neer test (NRS score ≥5) during the HBN maneuver and a difference of more than 10° of passive range of external rotation between shoulders had a 98% probability of predicting the presence of hemiplegic shoulder pain (receiver operating characteristic, 994; sensitivity, 96.7%; specificity, 99.0%; positive predictive value, 96.7%; negative predictive value, 99.0%; P<.001).

Conclusions: Three diagnostic clinical tests that can be performed during a bedside evaluation increase the likelihood of determining those who complain of hemiplegic shoulder pain after an acute episode of stroke.

Key Words: Hemiplegia; Musculoskeletal system; Pain; Rehabilitation; Shoulder pain.

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

Hemiplegic Shoulder Pain is among the 4 most common, yet preventable, medical complications that patients may experience after stroke.1-5 It can commence as early as the first 2 weeks after stroke.1,3-7 Seventeen percent of patients experience shoulder pain within the first week,3 55% within 2 weeks,2,5,7,12-17 87% by 2 months,2 and 75% within the first year after stroke.6-10 The early onset of hemiplegic shoulder pain adversely hindered the recovery of upper-limb function because it has been reported that improvement in upper-limb function within the first 5 weeks poststroke translated to greater usage of the affected limb during performance of functional tasks.11

Musculoskeletal conditions of noncentral origin such as glenohumeral instability, rotator cuff tears, subacromial and supraspinatus impingement, shoulder-hand syndrome, biceps tendonitis, and shoulder muscle atrophy also occurred after stroke.7,8,12-17 Less than 2 months after stroke, arthroscopic evaluations by Lo et al13 identified 11 different types of shoulder pain etiology. Fifty percent of patients experienced adhesive capsulitis13 and the incidence rose to 74% within a year.18 Restriction in shoulder external rotation also occurred among patients with severe shoulder pain less than 3 months after stroke.18-20 Some authors attributed the early limitation in shoulder ranges of motion (ROMs) to soft tissue contractures15 because patients with hemiplegic shoulder pain showed synovial hypervascularization and cellular proliferation without inflammatory infiltration.20 Other histologic findings among patients with adhesive capsulitis and shoulder-hand syndrome included elevated growth factor β, fibroblasts, tumor necrosis factor α, and infiltration of perivascular leukocytes into the capsule.22,23 Capsular adhesions further limited passive shoulder external rotation.22,24

During the early phase of recovery, 65% of patients also experienced muscle weakness with no or minimal degenerative changes within the glenohumeral joint.4 Patients with both low muscle tone and hemiplegic shoulder pain were at a greater risk of experiencing glenohumeral instability and impingement syndromes.7,14,22,25,26 Glenohumeral instability and subacromial impingement are commonly diagnosed by simple and noninvasive apprehension, Neer, and sulcus tests,27 but these tests are not usually performed on patients early after the onset of their stroke. Furthermore, there is a lack of reliable diagnostic clinical tests to identify the exact musculoskeletal etiologies associated with hemiplegic shoulder pain.

The purpose of this study was to explore the association between early reporting of hemiplegic shoulder pain among patients who experienced an acute episode of stroke with differences in their shoulder ROMs, muscle strengths, muscle tones, and positive findings of musculoskeletal clinical tests to predict the likelihood of underlying shoulder dysfunction. The secondary aim of the study was to establish valid and simple diagnostic clinical tests that could predict those patients at risk of developing hemiplegic shoulder pain.
METHODS

Design
The institutional review boards of 3 general hospitals approved this multicenter study that was conducted between October 2002 and May 2003.

Participants
Physical therapists from the hospitals recruited 152 adult inpatients who had experienced a first episode of acute stroke. Stroke was diagnosed as a focal disturbance of cerebral function due to vascular origin lasting more than 24 hours. Patients were included in the study if they were medically stable, cognitively oriented, and able to verbally express their needs. The exclusion criteria were bilateral stroke, thalamic infarcts, upper-limb sensory deficit, previous shoulder injury or pathology, reflex sympathetic dystrophy, receptive dysphasia, or unstable angina. The cuff tightness test described by Price et al28 confirmed that patients had the ability to differentiate levels of pain intensities. Seventeen patients were rejected because they showed mild alterations in their arm sensation or had early signs of shoulder-hand syndrome. One hundred thirty-five patients provided informed consent before they participated in the study.

Measurements
The definition of shoulder pain was in accordance with the Arthritis and Rheumatism Council Epidemiological Research Unit criteria of an expression of continuous pain.29 The 11-point numeric rating scale (NRS) quantified the intensity of pain, with 10 being the most intense level of pain and 0 being no pain. The NRS is a valid and sensitive graded pain intensity scale that is simple to use and highly recommended for studies involving older patients.30 In the present study, 30 patients reported moderate and greater (NRS score ≥5) intensities of shoulder pain at rest during their hospital stay. They were the experimental group, referred to as hemiplegic shoulder pain in the present study. The rest of the recruited patients (controls) reported mild or no shoulder pain at rest.

Physical therapists conducted all clinical tests at the patients’ bedside or during rehabilitation sessions within their hospital stay (mean, 8.62 ± 6.6d). These tests included the following.

Clinical examination of shoulder ROMs and muscle strengths. A goniometer was used to quantify the pain-free ranges of passive flexion, passive abduction, and passive external rotation of all patients. Measurements were taken while patients were supine or side-lying and in accordance with the established protocol. ROMs were compared with the unaffected upper limb and the differences expressed as a percentage of the available range of the unaffected upper limb in a particular direction. Passive ROM of all shoulder motions was divided into 2 categorical variables, that is, 10° of difference or less and more than 10° of difference.

The ordinal scale of the manual muscle test quantified muscle strength of the anterior, middle, and posterior fibers of the deltoid muscles. The scale reported a very good to good interrater (weighted κ range, .85—.94) and intrarater (weighted κ range, .81—.96) agreements when performed on patients who experienced an acute episode of stroke.31,32 The muscle strength of the deltoids were categorized into less than gravity, and equal to or greater than gravity. To keep the assessment simple and quick to perform, the strength of other shoulder muscles was not tested.

The Modified Ashworth Scale (MAS) quantified upper-limb muscle tone, because the scale has been reported to have good to very good interrater (weighted κ range, .77—.96) and intrarater (weighted κ range, .77—.83) agreements among patients who experienced stroke.33 Muscle tone of the upper limb was categorized as normal when the MAS score was 0 to 1+ and high when the score was 2 to 4. Low tone reflected a flaccid upper limb.

Combined upper-limb movements. The performance of dorsum of hand to lumbosacral junction (hand-behind-back [HBB]) maneuver reflects the combination of shoulder internal rotation and extension, whereas hand-behind-neck (HBN) maneuver is a combination of shoulder external rotation and abduction.34 The therapist placed each patient’s affected arm passively in both positions, one at a time, while patients reported the intensity of shoulder pain they experienced.

Musculoskeletal tests. Therapists performed the apprehension, Neer, and sulcus tests on the affected upper limb of all patients. The apprehension test is performed by placing the patient in supine position with their arm externally rotated, in abduction and slight extension. Reporting of shoulder pain or signs of apprehension during the test suggest the likelihood that the patient may have signs of anterior shoulder instability (sensitivity, 63%).27,35-37 In the sitting position, the therapist performed the Neer test by limiting each patient’s scapular rotation while internally rotating the affected arm in a passive mode through elevation in the scaphoid plane.27,30 Shoulder pain during this test is suggestive of subacromial impingement.27,35-38 The Neer test has a sensitivity of 88.7%.27,35-38 The sulcus test was performed in the sitting position with patient’s affected arm by their side while the therapist pulled the elbow inferiorly to measure the physiologic separation between the acromial and the humeral head. Separation of less than 1cm is scored as 1, 1 to 2cm is scored as 2, and greater than 2cm is scored as 3. A grade 3 sulcus test indicates likely etiology of multidirectional glenohumeral instability. The test has a sensitivity of 28%.27,35,36

Reliability Phase
A reliability study evaluated the agreement of recording of hemiplegic shoulder pain, HBB maneuver, HBN maneuver, posterior deltoid strength, and passive external rotation performed on 17 patients by 2 physical therapists. The second physical therapist was unaware of the results of the first therapist and revisited all 17 patients at least 2 hours after the completion of the initial examination. There was a high degree of reliability between the 2 therapists (table 1).

Statistical Analysis
Reports indicated that approximately 20% of patients experienced hemiplegic shoulder pain during their hospital stay.5,17,24 Thus, the minimum sample size required was identified to be

<table>
<thead>
<tr>
<th>Table 1: Interrer Reliability of Selected Clinical Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical Tests</td>
</tr>
<tr>
<td>Complaint of hemiplegic shoulder pain at rest</td>
</tr>
<tr>
<td>HBN maneuver</td>
</tr>
<tr>
<td>HBB maneuver</td>
</tr>
<tr>
<td>Difference of passive ROM of external rotation</td>
</tr>
<tr>
<td>Strength of posterior deltoid</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; ICC, intraclass confidence coefficient.
132 patients at a 95% confidence interval (CI) (15–25). SPSS® for Windows facilitated statistical analysis. During the reliability phase of this study, intraclass correlation coefficient (ICC) at 95% CI was calculated. Parametric t tests determined mean differences between the hemiplegic shoulder pain and controls. Chi-square tests evaluated the association between patients with hemiplegic shoulder pain and the categorical variables. Clinical measures of HBN maneuver, HBB maneuver, passive shoulder ROMs, and muscle strengths were categorized to obtain receiver operating characteristic (ROC) values from binary logistic regression analyses. The full predictive clinical model consisted of the highest significant ROC value for each clinical test. This study also reported sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive and negative diagnostic likelihood ratios (LR+; LR–), respectively, and probability of all clinical tests and variables. Significance level of all tests was \( P < 0.05 \).

RESULTS

Differences and Associations of Hemiplegic Shoulder Pain

Table 2 summarizes the characteristics and stroke status of all 135 patients. Most patients (83%) were aged over 55 years with a mean age of 64.36 ± 10.82 years. Exactly 81.5% experienced an infarction that led to a stroke, which was consistent with the established epidemiologic health trends of the local population. In our study, 22.2% (95% CI, 15.5–30.2) of patients reported hemiplegic shoulder pain at rest (NRS score ≥5). There was no association between hemiplegic shoulder pain at rest with age, sex, type of stroke, side of stroke, or muscle tone (\( P > 0.05 \)). All clinical examinations of shoulder ROMs and shoulder muscle strengths differed statistically between hemiplegic shoulder pain and control patients (\( P < 0.001 \)). Hemiplegic shoulder pain at rest was associated with lesser deltoid muscle strengths (\( P < 0.001 \)) and greater than 10° differences in passive flexion (\( P = 0.01 \)), passive abduction (\( P < 0.001 \)), and passive external rotation (\( P < 0.001 \)). However, there was no association between hemiplegic shoulder pain and passive ROM extension (\( P = 0.275 \)).

Predictive Clinical Tests of Hemiplegic Shoulder Pain

Table 3 shows diagnostic properties of individual clinical tests and their cutoff based on the highest ROC values (\( P < 0.05 \)). The complete predictive model consisted of positive apprehension and Neer tests, HBN maneuver, HBB maneuver, passive flexion of 5 or greater, passive flexion difference of more than 5°, passive extension difference of more than 10°, passive abduction difference of more than 5°, passive external rotation difference of more than 10°, and muscle strength in the 3 deltoid muscles of less than gravity. The ROC of the complete model of 11 clinical tests was 0.999 and the ROC remained unchanged when the predictive model consisted of only 6 clinical tests (table 4, model A). The ROC values of predictive models B and C consisting of 4 and 3 clinical tests, respectively, were slightly less. However, their sensitivity, specificity, PPV, NPV, LR+ and LR– were the same as model A. Model D versions I through III were combinations of 2 clinical tests each and reported smaller ROC values compared with model C. Although model C required only 3 clinical tests, when all 3 tests were positive there was a 98% probability of identifying complaints of hemiplegic shoulder pain at rest (table 5).

DISCUSSION

The current study found that the proposed predictive model consisting of 3 simple bedside clinical tests allowed clinicians

<table>
<thead>
<tr>
<th>TABLE 2: Characteristics, Stroke Status, and Clinical Examinations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variables</td>
</tr>
<tr>
<td>-----------</td>
</tr>
<tr>
<td>Mean age ± SD (y)</td>
</tr>
<tr>
<td>Age distribution (y)</td>
</tr>
<tr>
<td>≤60</td>
</tr>
<tr>
<td>61–70</td>
</tr>
<tr>
<td>≥71</td>
</tr>
<tr>
<td>Sex</td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Female</td>
</tr>
<tr>
<td>Type of stroke</td>
</tr>
<tr>
<td>Hemorrhage</td>
</tr>
<tr>
<td>Infarction</td>
</tr>
<tr>
<td>Side of stroke</td>
</tr>
<tr>
<td>Right</td>
</tr>
<tr>
<td>Left</td>
</tr>
<tr>
<td>Upper-limb muscle tone</td>
</tr>
<tr>
<td>Low (flaccid upper limb)</td>
</tr>
<tr>
<td>Normal (0 to 1+)</td>
</tr>
<tr>
<td>High (2 to 4)</td>
</tr>
<tr>
<td>Anterior deltoid strength</td>
</tr>
<tr>
<td>&lt; gravity</td>
</tr>
<tr>
<td>≥ gravity</td>
</tr>
<tr>
<td>Middle deltoid strength</td>
</tr>
<tr>
<td>&lt; gravity</td>
</tr>
<tr>
<td>≥ gravity</td>
</tr>
<tr>
<td>Posterior deltoid strength</td>
</tr>
<tr>
<td>&lt; gravity</td>
</tr>
<tr>
<td>≥ gravity</td>
</tr>
<tr>
<td>Passive ROM flexion</td>
</tr>
<tr>
<td>10° difference</td>
</tr>
<tr>
<td>&gt;10° difference</td>
</tr>
<tr>
<td>Passive ROM extension</td>
</tr>
<tr>
<td>10° difference</td>
</tr>
<tr>
<td>&gt;10° difference</td>
</tr>
<tr>
<td>Passive abduction</td>
</tr>
<tr>
<td>10° difference</td>
</tr>
<tr>
<td>&gt;10° difference</td>
</tr>
<tr>
<td>Passive external rotation</td>
</tr>
<tr>
<td>10° difference</td>
</tr>
<tr>
<td>&gt;10° difference</td>
</tr>
</tbody>
</table>

NOTE. Values are percent unless otherwise indicated. Abbreviation: SD, standard deviation. *Significant at \( P < 0.01 \).
Our results confirmed reports of an association between shoulder pain at rest after stroke and decreased shoulder external rotation in the affected shoulder. Limitation of shoulder external rotation on the paretic upper limb also correlated with the time of onset of stroke \( r = -0.538, P < 0.01 \). The current finding indicated that as early as a week after stroke, 93.3% of patients with hemiplegic shoulder pain at rest had a difference of more than 10° in range of shoulder external rotation between limbs. Although obligatory external rotation is necessary for pain-free gliding of the supraspinatus.

### Table 3: Diagnostic Properties of Significant Clinical Tests Among Patients Who Complained of Hemiplegic Shoulder Pain at Rest (n=30)

<table>
<thead>
<tr>
<th>Tests</th>
<th>ROC</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
<th>LR+</th>
<th>LR−</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apprehension test (+ve)</td>
<td>.695</td>
<td>80.0</td>
<td>59.0</td>
<td>35.8</td>
<td>91.2</td>
<td>2.0</td>
<td>3.0</td>
<td>.001</td>
</tr>
<tr>
<td>Neer test (+ve)</td>
<td>.881</td>
<td>100.0</td>
<td>76.2</td>
<td>54.5</td>
<td>100.0</td>
<td>4.2</td>
<td>∞</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>HBN maneuver (NRS)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥2</td>
<td>.648</td>
<td>100.0</td>
<td>29.5</td>
<td>28.8</td>
<td>100.0</td>
<td>1.4</td>
<td>∞</td>
<td>.014</td>
</tr>
<tr>
<td>≥3</td>
<td>.848</td>
<td>100.0</td>
<td>69.5</td>
<td>48.4</td>
<td>100.0</td>
<td>3.3</td>
<td>∞</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>≥4</td>
<td>.848</td>
<td>100.0</td>
<td>69.5</td>
<td>48.4</td>
<td>100.0</td>
<td>3.3</td>
<td>∞</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>≥5*</td>
<td>.979</td>
<td>96.7</td>
<td>99.0</td>
<td>96.7</td>
<td>30.0</td>
<td>2.3</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td>≥6</td>
<td>.783</td>
<td>56.7</td>
<td>100.0</td>
<td>89.0</td>
<td>∞</td>
<td>2.3</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td>≥7</td>
<td>.683</td>
<td>36.7</td>
<td>100.0</td>
<td>84.7</td>
<td>∞</td>
<td>1.6</td>
<td>&lt;.002</td>
<td></td>
</tr>
<tr>
<td>HBB maneuver (NRS)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥3</td>
<td>.652</td>
<td>100.0</td>
<td>30.5</td>
<td>29.1</td>
<td>100.0</td>
<td>1.4</td>
<td>∞</td>
<td>.011</td>
</tr>
<tr>
<td>≥4</td>
<td>.738</td>
<td>100.0</td>
<td>47.6</td>
<td>35.3</td>
<td>100.0</td>
<td>1.9</td>
<td>∞</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>≥5*</td>
<td>.845</td>
<td>96.7</td>
<td>72.4</td>
<td>50.0</td>
<td>98.7</td>
<td>3.5</td>
<td>21.9</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>≥6</td>
<td>.748</td>
<td>66.7</td>
<td>82.9</td>
<td>52.6</td>
<td>89.7</td>
<td>3.9</td>
<td>2.5</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>≥7</td>
<td>.767</td>
<td>53.3</td>
<td>100.0</td>
<td>88.2</td>
<td>∞</td>
<td>2.1</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td>≥8</td>
<td>.700</td>
<td>16.7</td>
<td>100.0</td>
<td>80.8</td>
<td>∞</td>
<td>1.2</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td>Passive flexion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;5° difference*</td>
<td>.781</td>
<td>93.3</td>
<td>62.9</td>
<td>41.8</td>
<td>97.1</td>
<td>2.5</td>
<td>9.4</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>&gt;10° difference</td>
<td>.719</td>
<td>60.0</td>
<td>83.8</td>
<td>51.4</td>
<td>88.0</td>
<td>3.7</td>
<td>2.1</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Passive extension</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;5° difference</td>
<td>.705</td>
<td>93.3</td>
<td>47.6</td>
<td>33.7</td>
<td>96.2</td>
<td>1.8</td>
<td>7.1</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>&gt;10° difference</td>
<td>.763</td>
<td>93.3</td>
<td>53.3</td>
<td>36.4</td>
<td>96.6</td>
<td>2.0</td>
<td>8.0</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>&gt;15° difference</td>
<td>.762</td>
<td>93.3</td>
<td>59.0</td>
<td>39.4</td>
<td>96.9</td>
<td>2.3</td>
<td>8.8</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>&gt;20° difference*</td>
<td>.781</td>
<td>93.3</td>
<td>62.9</td>
<td>41.8</td>
<td>97.1</td>
<td>2.5</td>
<td>9.4</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>&gt;25° difference*</td>
<td>.757</td>
<td>73.3</td>
<td>78.1</td>
<td>48.9</td>
<td>91.1</td>
<td>3.3</td>
<td>2.9</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Passive abduction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;5° difference*</td>
<td>.867</td>
<td>93.3</td>
<td>80.0</td>
<td>57.1</td>
<td>97.7</td>
<td>4.7</td>
<td>11.9</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>&gt;10° difference</td>
<td>.686</td>
<td>46.7</td>
<td>90.5</td>
<td>58.3</td>
<td>85.6</td>
<td>4.9</td>
<td>1.7</td>
<td>&lt;.002</td>
</tr>
<tr>
<td>Passive external rotation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;5° difference</td>
<td>.833</td>
<td>93.3</td>
<td>73.3</td>
<td>50.0</td>
<td>97.5</td>
<td>3.5</td>
<td>10.9</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>&gt;10° difference*</td>
<td>.881</td>
<td>93.3</td>
<td>82.9</td>
<td>60.9</td>
<td>97.8</td>
<td>5.5</td>
<td>12.4</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>&gt;15° difference</td>
<td>.790</td>
<td>66.7</td>
<td>91.4</td>
<td>69.0</td>
<td>90.6</td>
<td>7.8</td>
<td>2.7</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Anterior deltoid &lt; gravity</td>
<td>.810</td>
<td>93.3</td>
<td>68.6</td>
<td>45.9</td>
<td>97.3</td>
<td>3.0</td>
<td>10.2</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Middle deltoid &lt; gravity</td>
<td>.793</td>
<td>96.7</td>
<td>61.9</td>
<td>42.0</td>
<td>98.5</td>
<td>2.5</td>
<td>18.8</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Posterior deltoid &lt; gravity</td>
<td>.795</td>
<td>93.3</td>
<td>65.7</td>
<td>43.8</td>
<td>97.2</td>
<td>2.7</td>
<td>9.8</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

Abbreviation: +ve = positive test value.
* Cutoff is highest ROC value.

### Table 4: Diagnostic Properties of Predictive Models of Clinical Tests for Early Identification of Patients Who Complained of Hemiplegic Shoulder Pain at Rest (n=30)

<table>
<thead>
<tr>
<th>Predictive Models</th>
<th>ROC</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
<th>LR+</th>
<th>LR−</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model A (ROC &gt;0.8)</td>
<td>.999</td>
<td>96.7</td>
<td>99.0</td>
<td>96.7</td>
<td>99.0</td>
<td>96.7</td>
<td>30.0</td>
<td></td>
</tr>
<tr>
<td>Neer (+ve), HBN maneuver ≥5, passive abduction &gt;5° difference, passive external rotation &gt;10° difference, anterior deltoid &lt; gravity</td>
<td>.994</td>
<td>96.7</td>
<td>99.0</td>
<td>96.7</td>
<td>99.0</td>
<td>96.7</td>
<td>30.0</td>
<td></td>
</tr>
<tr>
<td>Model B (ROC &gt;0.85)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neer (+ve), HBN maneuver ≥5, passive abduction &gt;5° difference, passive external rotation &gt;10° difference</td>
<td>.994</td>
<td>96.7</td>
<td>99.0</td>
<td>96.7</td>
<td>99.0</td>
<td>96.7</td>
<td>30.0</td>
<td></td>
</tr>
<tr>
<td>Model C (ROC &gt;0.88)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neer (+ve), HBN maneuver ≥5, passive external rotation &gt;10° diff</td>
<td>.994</td>
<td>96.7</td>
<td>99.0</td>
<td>96.7</td>
<td>99.0</td>
<td>96.7</td>
<td>30.0</td>
<td></td>
</tr>
<tr>
<td>Model D (2 clinical tests, 3 versions)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(I): Neer (+ve) and HBN maneuver ≥5</td>
<td>.991</td>
<td>96.7</td>
<td>99.0</td>
<td>96.7</td>
<td>99.0</td>
<td>96.7</td>
<td>30.0</td>
<td></td>
</tr>
<tr>
<td>(II): Neer (+ve) and passive external rotation &gt;10° diff</td>
<td>.963</td>
<td>82.4</td>
<td>98.0</td>
<td>93.3</td>
<td>94.3</td>
<td>41.2</td>
<td>5.6</td>
<td></td>
</tr>
<tr>
<td>(III): HBN maneuver &gt;4 and passive external rotation &gt;10° difference</td>
<td>.992</td>
<td>96.7</td>
<td>99.0</td>
<td>96.7</td>
<td>99.0</td>
<td>96.7</td>
<td>30.0</td>
<td></td>
</tr>
</tbody>
</table>
tendon during arm elevation, 17% of patients who did not complain of hemiplegic shoulder pain also had limitation in shoulder external rotation on their paretic limb (see table 2). Patients with shoulder instability also reported signs of shoulder impingement during arm elevation. More investigation is required to evaluate the dynamics that occur at the glenohumeral joint of patients with hemiparetic upper limbs when they perform elevation tasks. Another study from our laboratory is currently studying the shoulder muscle activation patterns among patients with anterior shoulder instability and those with hemiparetic shoulder dysfunction to better understand the neuromotor control strategies that patients adopt during pain-free performance of overhead reaching tasks.

A number of researchers have questioned the accuracy of reporting pain associated with this poorly understood phenomenon of hemiplegic shoulder pain. Price et al expressed that stroke subtypes influenced the reporting of pain intensity at the shoulder. However, Bohannon and Andrews felt that the method of assessing shoulder pain was the most important factor. Constant and Murley assessed shoulder pain intensity among orthopedic patients during the performance of functional tasks. In the current study, shoulder pain was also assessed during performance of functional tasks including HBB and HBN maneuvers, and the cutoff intensity level of reporting of shoulder pain was moderate (NRS score ≥5), which was consistent with findings of another study involving 54 patients. They too reported that approximately 20% of stroke patients experienced shoulder pain at intensities equal or greater than 5. Currently, there is no reliable criterion standard to accurately evaluate signs of acute shoulder pain. The present study found that complaints of moderate pain at rest and 3 positive clinical tests can act as a pseudo-valid standard to identify those who are at risk of hemiplegic shoulder pain. Moreover, the physical therapist’s interrater reliability measurements of shoulder external rotation in the present study (see table 1) were similar to scores reported by Braus et al. Our ICC values were also better than .19 and .73 scored by rheumatologists and physical therapists, respectively, when they measured shoulder external rotation during abduction with goniometry (see table 1). An ICC value of .89 is deemed to be acceptable for clinical practice.

Chronic conditions such as complex regional pain syndrome have identified predictive symptoms and their critical levels, and these findings led to better understanding and management of this chronic condition. To our knowledge, this study is the first to establish that a combination of 3 simple clinical tests can accurately predict the early onset of shoulder pain after stroke. The results enrich our understanding of the possible pathophysiology associated with this poorly understood yet preventable medical condition. We plan to further evaluate the

### Table 5: Diagnostic Probabilities of Predictive Model C That Can Be Used to Identify Patients With Complaints of Hemiplegic Shoulder Pain at Rest (n=30)

<table>
<thead>
<tr>
<th>Neer Test</th>
<th>HBB Maneuver</th>
<th>Passive External Rotation Difference (deg)</th>
<th>Probability (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative</td>
<td>≤4</td>
<td>≤10</td>
<td>0.0</td>
</tr>
<tr>
<td>Negative</td>
<td>≥5</td>
<td>≤10</td>
<td>0.0</td>
</tr>
<tr>
<td>Positive</td>
<td>≥5</td>
<td>&gt;10</td>
<td>1.7</td>
</tr>
<tr>
<td>Positive</td>
<td>≤4</td>
<td>&gt;10</td>
<td>11.0</td>
</tr>
<tr>
<td>Positive</td>
<td>≥5</td>
<td>&gt;10</td>
<td>84.0</td>
</tr>
<tr>
<td>Positive</td>
<td>≥5</td>
<td>&gt;10</td>
<td>98.0</td>
</tr>
</tbody>
</table>

3 musculoskeletal tests identified against examination under anesthesia and magnetic resonance imaging of the shoulder, which are deemed to be the criterion standards to evaluate shoulder dysfunction.

### Study Limitations

This study quantified pain intensity at the hemiplegic shoulder using an NRS, which is easy to administer and score. We adopted this approach because there is no criterion standard to predict the onset of hemiplegic shoulder pain. To minimize false-positive reporting of shoulder pain, the current study raised the cutoff intensity threshold to an NRS score of 5 or more, consistent with practices of others. Furthermore, the cuff tightness test verified the ability of patients to discriminate the intensity of pain in the upper limbs. The current study could have used another pain intensity measuring tool to confirm reliability of the NRS, but this would have prolonged the assessment procedure and probably caused frustrated patients to respond inaccurately. We acknowledge that it would be ideal to perform costly investigations such as radiography, magnetic resonance imaging, or ultrasound as criterion standards to evaluate all 135 patients participating in the current study. With the results of the current study, we can now incorporate these investigative procedures for a smaller population of acute stroke patients to evaluate the concurrent validity of the current battery of 3 clinical tests to predict the early onset of hemiplegic shoulder pain. Another possible limitation of the current study was the inclusion of patients who suffer infarctions at various sites of the brain. Although our exclusion criterion eliminated patients with thalamic infarctions and central shoulder pain syndrome, neural reorganization after stroke may have stimulated the thalamus to induce shoulder pain. The unaffected ipsilateral shoulder may also have impaired ranges of shoulder motion and muscle strengths that could provide inaccurate comparisons. Finally, there were only 30 patients in the experimental group who met the criteria of moderate hemiplegic shoulder pain. The small number of patients in the experimental group suggests that the current predictive model requires further investigation for universal acceptance. Thus, we are embarking on recruiting more patients to supplement the current study’s data pool.

### CONCLUSIONS

The study provides 3 simple clinical tests to predict those who are likely to complain of hemiplegic shoulder pain early after stroke. Early diagnosis of this pathology will encourage early intervention and presumably an improved outcome for the patients.

### Acknowledgments

We thank the physicians and therapists from Tan Tock Seng Hospital, Changi General Hospital, and National University Hospital for their active support during the course of this study.

### References

41. Marx RG, Bombardier C, Wright JG. What do we know about the reliability and validity of physical examination tests used to examine the upper extremity? J Hand Surg [Am] 1999;24:185-93.

Objective: To investigate the variation of muscle coactivation patterns during the course of robot-assisted rehabilitation on elbow flexion and extension for chronic stroke.

Design: A detailed electromyographic analysis was conducted on muscle activation levels and muscle coactivation patterns, represented by a cocontraction index of a muscle pair, for the muscles of biceps brachii, triceps brachii, anterior deltoid, and posterior deltoid, during training of elbow extension and flexion, actively assisted by a robot, from 0° to 90° by tracking a target moving at a speed of 10°/s on the screen.

Setting: Rehabilitation center research laboratory.

Participants: Seven hemiplegic chronic stroke patients received elbow training.

Interventions: Each subject received 20 sessions (1.5 hours/session) of the elbow training on his/her paretic side at an intensity of 3 to 5 times a week for a training period of 7 consecutive weeks.

Main Outcome Measures: Muscle cocontraction index, muscle activation level, and Modified Ashworth Scale (MAS), Fugl-Meyer Assessment (FMA), and Motor Status Scale (MSS) scores.

Results: The electromyographic activation levels of the biceps brachii, triceps brachii, and anterior deltoid of each subject decreased during the training. The overall electromyographic activation levels of the biceps and triceps, which summarized the performance of all subjects, decreased significantly in the middle sessions (from the 8th to 12th sessions) of the training (P<.05), associated with the significant decrease (P<.05) in the MAS score. The overall electromyographic activation level of the anterior deltoid also decreased significantly from the 8th to 20th sessions (P<.05). Significant decreases in the cocontractions of all muscle pairs were observed in all subjects and also in the overall cocontraction index (P<.05). The cocontraction between the biceps and triceps significantly decreased when the overall electromyographic levels of the 2 muscles were stable from the 10th to 20th sessions (P<.05). Significant improvements (P<.05) on the FMA and MSS score were also found by the pre- and postassessments.

Conclusions: In the 20-session robot-assisted training, the excessive muscle activations reduced mainly in the first half of the training course, which could be related to the learning process of the tracking skill and also to the reduction in muscle spasticity. The muscle coordination for achieving elbow tracking improved significantly in the latter sessions of the training, represented as decreased cocontraction indexes between the muscle pairs.

Key Words: Assistive technology; Electromyography; Muscles; Rehabilitation; Robotics; Stroke.

From the Department of Health Technology and Informatics, Hong Kong Polytechnic University, Hong Kong.

Supported by the Research Grants Council of the Hong Kong Special Administrative Region, China (grant no. PolyU 527105E) and The Hong Kong Polytechnic University (grant nos. G-U056, G-YX65).

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.

Reprint requests to Kai Y. Tong, PhD, Dept of Health Technology and Informatics, The Hong Kong Polytechnic University, Rm ST417, Core S, 4/F, Kowloon, Hong Kong, e-mail: k.y.tong@polyu.edu.hk. 0003-9993/07/8808-1147/4$32.00/0 © 2007 by the American Congress of Rehabilitation Medicine and the American Academy of Physical Medicine and Rehabilitation.
ments were more effective for motor functional improvement than treatments using continuous passive motions. Robotic treatment with active-resisted motion involved voluntarily completing movements against programmed resistance. It has been found that repetitive practice of hand and finger movements against loads resulted in greater improvements in motor performance and functional scales than Bobath-based treatment. Transcutaneous electric nerve stimulation, and supra-threshold electric stimulation on hand and wrist muscles. In robotic therapy, repeated practice against opposing force mainly improved the muscle force from the elbow and shoulder in a reaching task, and also benefited the functional improvements in the wrist and hand for chronic stroke.

Despite positive documentation of overall clinical outcomes after robot-assisted rehabilitation for chronic stroke, the precise effects of the interventions on the motor system recovery have not been well described. A solution to this is to follow the evolution of specific markers of motor ability over the course of rehabilitation treatment. Changes of movement smoothness during robot-assisted stroke recovery have been described by Rohrer et al; however, the kinematic parameters used in their study for the evaluation of movement smoothness did not directly reveal the evolution of poststroke motor system during rehabilitation. Impairment in hemiparetic stroke is usually accompanied by abnormalities of spasticity, muscle weakness, and disturbances in muscular coordination mainly reflected by varied muscle coactivation patterns. In many hemiparetic persons after stroke, when the physical signs (spasticity, weakness) have been treated effectively, motor impairment associated with abnormal muscle coactivation patterns could still be present and severe. Dewald et al. suggested that the primary source of motor dysfunction or global disability in many hemiparetic patients after stroke was abnormal movement coordination, that is, abnormal muscle coactivation pattern. With the quantitative analyses on the electromyography recorded from the paretic upper limb in persons after stroke, reduced muscle coactivation patterns were found, and a relatively high correlation of motion at adjacent joints was also observed during isometric contractions in the previous works conducted by Dewald and colleagues. The related clinical observation is that the muscle coordination in hemiparetic persons after stroke is almost entirely stereotyped and does not permit different combinations of muscles. To our knowledge, muscle coactivation patterns in chronic stroke patients over the course of robot-assisted training have not been thoroughly studied yet. The purpose of the current study was to quantitatively investigate the recovery process in motor control related to the elbow and shoulder joints in chronic stroke by monitoring the evolution of muscle coactivation patterns during an elbow flexion and extension training actively assisted by a rehabilitation robot.

METHODS

After obtaining approval from the Human Subjects Ethics Sub-Committee of the Hong Kong Polytechnic University, we recruited 7 hemiplegic subjects after stroke for the study. All of the subjects were in the chronic stage (at least 1 year postonset of stroke; 6 men, 1 woman; age, 51.1 ±9.7y). All subjects received a robot-assisted elbow training program consisting of 20 sessions, with at least 3 sessions a week and at most 5 sessions a week, and finished in 7 consecutive weeks. Each training session was completed in 1.5 hours. Before and after the training, we adopted 2 clinical scales to evaluate the voluntary motor function of the paretic upper limb (the elbow and shoulder) of the subjects: the Fugl-Meyer Assessment (FMA; for elbow and shoulder; maximum score, 42) and the Motor Status Scale (MSS; shoulder/elbow; maximum score, 40). Spasticity of the paretic elbow of each subject before and after the training was assessed by the Modified Ashworth Scale (MAS) score. The clinical assessments of this study were conducted by a blinded therapist.

During each training session, each subject was comfortably seated, and the affected upper limb was placed horizontally on an electromyography-driven motor system (a Dynaserv motor, associated with an AKC-205A torque sensor; accuracy, 0.03Nm) developed by Song et al. with the elbow joint positioned at the origin, as shown in figure 1. The forearm of the affected side was placed on a manipulandum, which could rotate with the motor; and the elbow angle signals were measured by the motor via readings of the positions of the manipulandum. A belt was used to fasten the shoulder joint in order to keep the joint position still during elbow extension and flexion. Electromyography electrode pairs with a center separation of 2cm were attached to the skin surface of the muscle belly of biceps brachii, triceps brachii (lateral head), anterior deltoid, and posterior deltoid, according to the configuration specified in Cram’s work. The electromyography electrode pairs were not moved once placed. The electromyographic signals were preamplified, band-pass filtered (from 10 to 500Hz) and recorded through an analog-to-digital card, together with the angle signals, with a sampling frequency of 1000Hz.

The electromyographic signals for the muscles of interest during the resting state were first recorded before any voluntary motion taken by a subject in each session, which served as the electromyographic baselines of the individual muscles for the session. The isometric maximum voluntary flexion (IMVF; duration, 5s) and extension (IMVE; duration, 5s) of the elbow at a 90° elbow angle were then measured at a repetition of 3 times, respectively, with a 5-minute rest break after each contraction to avoid muscle fatigue. Other studies have reported that the maximum extension and flexion torques could usually be found at 90° of the elbow angle for hemiplegic subjects after stroke. During the training, each subject was required to carry out voluntary elbow flexion and extension in the elbow range from 0° to 90° (0° representing full extension) by tracking a target cursor moving at an angular velocity of 10° per second on the screen for both flexion and extension. From the study by Cheng et al and our own experimental experiences, 10 per second was chosen as a reasonable speed for subjects after stroke to follow, in order to prevent too difficult or too easy a pace for the subjects to achieve. Background music and verbal encouragement were given to subjects during the training. During tracking, active-assisted torques were generated by the motor system during the extension only, based on the robot design, with the supportive torque controlled by electromyographic signals. The active-assisted torque during the extension movement was defined as:

\[
T_{e} = G \cdot T_{IMVE} \cdot M_{e}
\]  

(1)

where G is a constant gain used to adjust the magnitude of the assistive torque and \(T_{IMVE}\) is the maximum value of the extension torque at the elbow angle of 90°. \(M_{e}\) in equation 1 was defined as

\[
M_{e} = \frac{EMG_{TRI} - EMG_{rest}}{EMG_{IMVE} - EMG_{rest}}
\]  

(2)

where \(EMG_{TRI}\) is triceps brachii electromyographic activity after the processes of full-wave rectification and moving average with a 100ms window, \(EMG_{rest}\) was the averaged \(EMG_{TRI}\) during the resting state, and \(EMG_{IMVE}\) was the maximum value of \(EMG_{TRI}\) during IMVE. The reasons for applying supportive torques in extension only were that hemiplegic
subjects usually have more difficulty in carrying out extension than flexion, and their flexors were commonly more spastic than extensors. It has been found that the elbow tracking and reaching performances of poststroke subjects could be immediately improved when using this type of active-assisted robot devices from different research group studies. In the current study, each subject was allowed to practice tracking for 10 minutes before the start of the training, to familiarize themselves with the course. In each training session, there were 18 tracking trials, and each trial had 5 cycles of elbow extension and flexion. In all trials, active-assisted torques were given in extension associated with the gain, G in equation 1, equal to 0%, 50%, and 100% alternatively applied to the tracking trials in a session. Resistive torques were also applied to each trial with values of a percentage of the torques during the maximum voluntary contractions (extension and flexion), that is,

$$T_r = a \cdot T_{MVC}$$

where $T_r$ was the resistive torque, $a$ was the percentage (10% or 20%, alternatively applied to the tracking trials in a session), and $T_{MVC}$ that included 2 parts, the maximum $T_{IMVF}$ (applied in the flexion phase only) and $T_{IMVE}$ (applied in the extension phase only). The net torque provided by the robot during the training is

$$T_n = T_s - T_r$$

where $T_n$ is the supportive torque and $T_r$ was the resistive torque. The purposes of applying the resistive torques proportional to the IMVF and IMVE during the training were (1) to improve the muscle force generation of the paretic limb, and (2) to keep the effective muscular effort at a level associated with a possible increase in muscle force during the training. Although $T_s$ and $T_r$ would tend to cancel, the 2 torques were directly related to the personal effort of the subjects during the training. Therefore, the net torque provided by the robot was interactive to the motor ability of subjects. Subjects were allowed to have a rest break of 2 minutes between consecutive tracking trials.

Electromyographic activity from the muscles of interest and angle signals during the training were recorded and stored in a computer during the even-numbered sessions of the training for offline processing. The elbow angle signals were low-pass filtered with a cutoff frequency of 20Hz. The torque signals during the IMVF and IMVE were also low-pass filtered with a cutoff frequency of 10Hz. A 4th-order, zero-phase forward and reverse Butterworth digital filter was adopted for the filtering processes. Figure 2A shows the representative signals recorded from a subject during the training. The linear envelope of the recorded electromyographic signals was obtained by (1) full-wave rectification, (2) low-pass filtering (10Hz cutoff frequency with fourth-order, zero-phase forward and reverse Butterworth filter), (3) subtraction of the average electromyographic activity during the resting state, and (4) normalized to the maximum value of electromyographic activation in each muscle during either a training session or the IMVF and IMVE of each session. Most of the electromyographic maximum values were observed in IMVF and IMVE, and only few (4 sessions) were found during the tracking task. The coactivations among muscle pairs during
the training were studied by the cocontraction index (CI) as introduced in Frost et al.’s study,\textsuperscript{27} that is,

$$\text{CI} = \frac{1}{T} \int_{0}^{T} A_{i,j}(t) \, dt$$

(3)

where $A_{i,j}(t)$ was the overlapping activity of electromyographic linear envelopes for muscles i and j, and $T$ was the length of the signal. The value of a cocontraction index for a muscle pair varied from 0 (no overlapping at all in the signal trial) to 1 (total overlapping of the 2 muscles with both electromyographic levels kept at 1 during the trial). The representative segments of electromyographic envelopes from the muscle pairs in a tracking trial are shown in figure 2B. The electromyographic activation level of a muscle in a tracking trial was also calculated by averaging the electromyographic envelope of the trial. The cocontraction indexes for different muscle pairs, the electromyographic activation levels of each muscle,
and the root mean square error (RMSE) between the target and the actual elbow angle were calculated for each trial of all even sessions. The averaged values of the cocontraction indexes and RMSEs of all trials in a session for each subject were used as the experimental readings for statistical analyses. Statistical analyses on the variation of the overall cocontraction indexes, overall electromyographic activation levels, and the overall RMSEs across the sessions, summarizing the performance of all subjects, were carried out by analyses of variance (ANOVAs) with Bonferroni post hoc test. A paired t test was used for comparison of the clinical scores before and after the training. The statistical significant level was .05.

RESULTS

Table 1 shows the clinical scales used for the impairment evaluation before and after 20-session training. It was found that the mean values of scores for FMA and MSS had increased significantly (t tests, P<.05) and the mean value of MAS decreased significantly (t tests, P<.05) after the training. For each subject, the clinical scales indicated improvements in at least 2 items.

Figure 3 shows the variation of the overall RMSE of the elbow angle during the tracking training. The overall RMSE varied significantly across the sessions with a decreasing tendency (1-way ANOVA, P<.05). The statistically significant decrease occurred at the 10th session compared with the values for the 2nd, 4th, and 6th sessions (post hoc tests, P<.05). The variation of RMSE from the 2nd to 8th sessions was not significant. There was also no significant changes in RMSE from the 10th to 20th sessions, except a local minimum at the 19th session (post hoc tests, P<.05). Decreasing tendencies in mean RMSE value were also observed in all individual subjects by comparing the mean RMSE values of the 2nd and 20th sessions and the decreases varied from 15.6% (subject 6) to 59% (subject 3). For subjects 1, 2, 3, 4, and 7, the maximum RMSEs were observed at the 2nd session, whereas for subjects 5 and 6, the maximum RMSEs appeared at the 6th session.

Figure 4 shows the electromyographic activation levels of each muscle during the training. The overall electromyographic activation level of the 4 muscles varied significantly across the sessions during the training (1-way ANOVA, P<.05). A significant decreasing tendency in the overall electromyographic activation level for the biceps brachii and triceps brachii mainly occurred before the 12th session (post hoc tests, P<.05). There was no significant decrease in the overall electromyographic activation level of the biceps brachii from the 10th session till the end, and the overall activation level of the triceps brachii did not change from the 12th to 20th sessions. The decreasing tendency of the overall anterior deltoid activation level was significant from the 8th session to the 20th session (post hoc tests, P<.05). There was no decrease found in the overall posterior deltoid activation level throughout the training; and the electromyographic levels at the 6th, 8th, and 20th sessions were even significantly higher than that of the 2nd session (post hoc tests, P<.05). The posterior deltoid activation levels for individual subjects did not display a consistent trend.

Figure 5 shows the muscle cocontraction patterns during the training, represented by the cocontraction index of each muscle pair. The variations in the overall cocontraction index of all muscle pairs were significant (1-way ANOVA, P<.05), and the overall cocontraction index of all muscle pairs reached their maximum at the 8th session. There was no significant change in the global RMSE summarizing the performance of all subjects during the tracking across the training sessions, represented by the values of mean and standard deviation (SD) (error bars).
found in the overall cocontraction indexes of the triceps brachii and posterior deltoid and biceps brachii and triceps brachii (see fig 5) during the first 8 sessions. The overall cocontraction indexes of the biceps brachii and anterior deltoid, anterior deltoid and posterior deltoid, and biceps brachii and posterior deltoid at the 8th session were significantly higher than the cocontraction indexes at the 2nd session (post hoc tests, $P<.05$). The overall cocontraction indexes of the muscle pairs biceps brachii and anterior deltoid, anterior deltoid and posterior deltoid, and triceps brachii and anterior deltoid reached a local minimum (post hoc tests, $P<.05$) at the 6th session before the appearance of the maximum mean values at the 8th session. For all muscle pairs, there was a significant decrease in the cocontraction index value from the 8th session to the 10th session (post hoc tests, $P<.05$). After the 8th session (from the 10th to 20th sessions), the overall cocontraction index values of the biceps brachii and triceps brachii, biceps brachii and anterior deltoid, anterior deltoid and posterior deltoid, and triceps brachii and anterior deltoid showed a significant decreasing tendency until the end of the training (post hoc tests, $P<.05$). The overall cocontraction indexes of the triceps brachii and posterior deltoid and biceps brachii and posterior deltoid varied nonsignificantly from the 10th session to the 20th session. The maximum values of the cocontraction indexes of the muscle pairs for each subject appeared mostly on or before the 8th session, except the cocontraction indexes of biceps brachii and anterior deltoid for subject 2 (at the 12th session), anterior deltoid and posterior deltoid for subject 2 (at the 10th session), triceps brachii and anterior deltoid for subject 4 (at the 16th session), and biceps brachii and posterior deltoid for subject 6 (at the 10th session). By comparing the maximum cocontraction index value with the cocontraction index at the last session, decreases in the cocontraction indexes of the muscle pairs for the individual subjects were found to vary from 7.6% (biceps brachii and posterior deltoid for subject 1) to 82.5% (biceps brachii and triceps brachii for subject 7).

**Fig 4.** The variations of electromyographic activation level for the biceps brachii, triceps brachii, anterior deltoid, and posterior deltoid muscles. The electromyographic activation levels from individual subjects (ie, S1 to S7) are represented by their respective symbols shown in the legend. The global electromyographic activation levels of the muscles, summarizing the performance of all subjects, are represented by values of mean and SD (error bar).

**Fig 5.** The variations in cocontraction index of the muscle pairs, triceps brachii and posterior deltoid, biceps brachii and triceps brachii, biceps brachii and anterior deltoid, anterior deltoid and posterior deltoid, triceps brachii and anterior deltoid, and biceps brachii and posterior deltoid. The cocontraction indexes of individual subjects are represented by symbols shown in the legend. The values of the global cocontraction indexes for a muscle pair summarizing the performance of all subjects are represented by the values of mean and SD (error bar).

**DISCUSSION**

After the 20 sessions of robot training of the elbow extension/flexion, motor improvements could be observed in all subjects, associated with the improved clinical scores, and decreases in the RMSE, cocontraction indexes, and electromyographic activation levels. The decrease in the clinical score of MAS suggested a reduction in spasticity of the impaired upper limb of the subjects. Increases in the FMA score and MSS implied improved motor functions of the paretic upper limb during prescribed voluntary movements. However, pre- and post-tests using the clinical scales can only provide general observations on the functional improvements, with little information revealed about the recovery process associated with the specific treatment.

The decrease in the mean values of the RMSE during tracking across the sessions in all subjects (see fig 3) suggested improvement in tracking performance. Adult cerebral cortex is capable of significant functional plasticity, and postinjury behavioral experience is a major modulator of neurophysiologic and neuroanatomic changes that take place in the undamaged tissue.28 It has been found that task-oriented physical training could help the redevelopment of lost motor functions by organizing neuromuscular pathways with compensatory motor centers.28 Intensive repeated robot-assisted training has helped the functional motor recovery process for persons after stroke even during the chronic state.28 Fasoli et al9 indicated that poststroke motor recovery was similar to motor learning to some extent, and what was known about motor learning may predict the course of motor recovery. In motor learning studies, the learning of a skilled movement has been characterized by a plateau of little or no change in performance.29 Therefore, in this work, after the 10th session, the tracking skill had been stably developed or learned by the subjects after stroke, because the RMSE reached its steady state after the 10th session.

The changes in muscle coactivation patterns during the task-oriented and robot-assisted training were analyzed by the elec-
tromyographic activation levels of individual muscles and the cocontraction indexes of different muscle pairs. Before the 10th session, the overall RMSE values during tracking were relatively high. This was associated with the higher overall electromyographic levels of the biceps brachii, triceps brachii, and anterior deltoid muscles in these sessions (see fig 4), during which the cocontractions, that is, the overall cocontraction indexes, between the different muscle pairs were also high (see fig 5). Two major reasons could explain the higher electromyographic levels before the 10th session: the overactivation of muscles during the initial period of motor learning for a skillful task,30 and the spasticity after stroke, which could cause extra muscle activities.31 The high cocontraction levels in the muscle pairs observed during this period were also mainly associated with the excessive electromyographic activities of the muscles. The significant decreases in the overall electromyographic levels of the biceps brachii and triceps brachii, that is, the main agonist and antagonist muscles related to the elbow joint, occurred in the middle sessions of the training (from the 8th to 12th sessions), but the overall electromyographic levels of the 2 muscles were almost stable in the latter sessions (after the 12th session). This possibly implies that the reduction in the muscle spasticity, measured by the MAS after the training, mainly occurred in the middle sessions (ie, from the 8th to 12th sessions). The reduction in the electromyographic levels between an antagonist muscle pair also released the cocontraction between the muscles during elbow flexion and extension, illustrated by the cocontraction index values in the latter content. It is also understood that the reduction in electromyographic activation level during tracking was possibly because of the increased muscle force with resistive training, that is, the subjects could perform the tracking tasks with less muscle effort. However, the resistance applied to each session was proportional to the maximum flexion and extension torques of the same session, which was associated with the possible improvement in muscle force generation. Therefore, the reductions in electromyographic levels observed should be mainly attributed to the decrease in excessive muscle activations. The decreasing trend of the overall anterior deltoid electromyographic level suggests the reduction in the excessive muscle activations of the shoulder during the elbow tracking. Compared with the anterior deltoid muscle, the overall electromyographic level of the posterior deltoid muscle decreased little across the training sessions. This is possibly because flexors are usually more spastic than extensors in persons after stroke and muscle weakness is more commonly observed in extensors rather than flexors.25,32

The cocontraction between the antagonist muscle pair around a joint in subjects without impairment could contribute to stabilizing of the joint in a static motion,33 and to movement accuracy in a dynamic motion.34 However, excessive cocontractions are energetically expensive,35 and abnormal muscle coactivation patterns, mainly cocontractions, in the paretic limb after stroke even degrade the accuracy and efficiency of limb movements.15 The significant decreases in the cocontraction index values of all muscle pairs (see fig 5) was associated with the improvement in the tracking accuracy represented by the reduced RMSE (see fig 3). These decreases suggest overall improvement in the coordination of the individual muscles, or more effective and efficient muscle coordination for achieving the tracking task by the elbow flexion and extension. There are 2 major reasons that could explain the decreased cocontraction index of a muscle pair: the reduction in the electromyographic activation level of the muscles, and the reduced cocontracting phase of the 2 muscles. The decrease in the overall cocontraction index of the triceps brachii and posterior deltoid was mainly due to the reduction in the electromyographic level of the triceps brachii, because the variation of this cocontraction index was consistent with the change in the overall triceps brachii electromyographic levels. However, for the overall cocontraction index between the biceps brachii and triceps brachii, further decrease after the 10th session was observed when the overall electromyographic levels of the 2 muscles were almost unchanged. This suggests that the decrease was not related to the reduction in the excessive electromyographic activation of the antagonist muscle pairs, but was associated with the reduced cocontracting phase of the biceps brachii and triceps brachii. It also indicated an improved contracting and relaxing phase pattern of the biceps brachii and triceps brachii during elbow extension and flexion in the latter sessions of the training. The decrease in the overall cocontraction index of the biceps brachii and triceps brachii appeared later, with a longer time course than the decreases in electromyographic levels of the respective biceps brachii and triceps brachii muscles. No steady state was reached by the decreasing cocontraction index of the biceps brachii and triceps brachii in the 20 sessions. It implied that further improvement in the muscle coordination between the biceps brachii and triceps brachii could possibly be obtained by providing additional sessions; and electromyographic cocontraction index, that is, cocontraction index, could be used as a simple measure for monitoring the improvement in muscle coordination during the training. For many persons after stroke, their elbow movements are usually associated with unnecessary shoulder activities. The decrease in the overall cocontraction indexes of the biceps brachii and anterior deltoid, triceps brachii and anterior deltoid, and biceps brachii and posterior deltoid after the 10th session indicated the better isolation of the elbow movement from the shoulder, which was related to the reduction in excessive activation of the anterior deltoid muscle. The decrease in the overall cocontraction indexes of the biceps brachii and anterior deltoid, anterior deltoid and posterior deltoid, and triceps brachii and anterior deltoid to the 20th session also suggests that further reduction in these cocontraction indexes was possible with more training sessions. It has been reported in existing literature that poststroke training for the distal joints could increase the motor capacity related to the intralimb proximal joint23 and the elbow training in this work might also benefit the functional recovery of the shoulder joint.

As reported by subjects in this study, the major improvement they all could feel after the training was the significant reduction in contracture and stiffness at the elbow and shoulder joints. This could be related to the great reduction in MAS score after the training (decreased almost half in the mean value compared with that of the pretraining) (see table 1). However, the improvements in FMA and MSS did not have as much change as in MAS after the training. Further improvement in motor functions assessed by FMA and MSS is possible with more training sessions in future studies, because the decrease in cocontraction index values did not reach a plateau after the 20 sessions of training in this study.

CONCLUSIONS

In this study, significant motor improvements assessed by MAS, FMA, and MSS were observed after the 20-session training on elbow tracking task actively assisted by a rehabilitation robot. The muscle coactivation patterns during the interactive robot-assisted training on elbow flexion and extension were analyzed by the electromyographic activation level of individual muscles and the electromyographic cocontraction index of the muscle pairs. The electromyographic activation levels of the major agonist and antagonist muscle pair of the elbow joint, biceps brachii and triceps brachii, significantly decreased in the first half of the training course, which was
associated with an improvement in tracking skill and a decrease in spasticity. The electromyographic level of the anterior deltoid also decreased during the training, suggesting a better isolation of elbow movements from the shoulder in the paretic limb. The coordination among the individual muscles related to the elbow and shoulder joints in the paretic upper limb improved mainly in the latter half of the training course, and was reflected in the reduction in the overall cocontraction indexes of the different muscle pairs. The results obtained in this study provide further understanding of the recovery process, especially muscle coordination, during interactive robot-assisted training, which would be useful for the design of robot-assisted training programs. We suggest that clinical studies with larger subject sample sizes to examine robot-assisted rehabilitation effects related to different training intensities could be carried out in the future studies.

References


Suppliers


b. Yungang West RD 17#, Feng Tai District, Beijing, China.

c. Noraxon dual electrodes; Noraxon USA Inc, 13430 N. Scottsdale Rd, Ste 104, Scottsdale, AZ 85243-1108.

d. INA126; Texas Instruments Inc, Product Information Center, 13532 N Central Expwy M/S 3807, Dallas, TX 75243-1108.

e. Noraxon dual electrodes; Noraxon USA Inc, 13430 N. Scottsdale Rd, Ste 104, Scottsdale, AZ 85243.

f. Yungang West RD 17#, Feng Tai District, Beijing, China.

Objectives: To determine (1) the frequency of osteoporosis at the hip and lumbar spine in a postpolio clinic population and (2) the association of lower-extremity muscle strength and other potential contributing factors to osteoporosis with bone density measured at the hip.

Design: Cross-sectional study involving a chart review.

Setting: A university-affiliated hospital postpolio clinic.

Participants: Patient charts (N=379) were reviewed; 164 (26%) were included, and 215 (74%) were not included primarily (74%) because of the unavailability of bone density results.

Interventions: Not applicable.

Main Outcome Measures: Bone density (in g/cm²) and T score were assessed at the femoral neck and lumbar spine. Muscle strength was evaluated by manual muscle testing in 7 bilateral lower-extremity muscles.

Results: The occurrence of osteoporosis at the hip and lumbar spine was 20 (32%) of 62 and 6 (10%) of 61 in men, 3 (9%) of 33 and 2 (6%) of 32 in premenopausal women, and 18 (27%) of 67 and 7 (11%) of 65 in postmenopausal women, respectively. In a logistic regression model, the presence of osteoporosis at the hip was significantly associated with strength sum score in the same extremity in which the bone density was performed after adjusting for other important risk factors (age, body mass index, time since polio).

Conclusions: Osteoporosis occurred commonly at the hip in a postpolio clinic population. Hip bone density was associated with muscle strength in the same lower extremity.

Key Words: Osteoporosis; Poliomyelitis; Rehabilitation; Risk factors.

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

...
A bone densitometry was requested depending on the doctor’s judgment after patient evaluation. A bone densitometry evaluation was considered for essentially all clinic patients. The examination was requested by the clinic physician if no such evaluation had been performed in the past few years in postmenopausal women and in premenopausal women and in men with lower-extremity weakness. All patients were evaluated by the same physician. The majority of the bone densitometry examinations were performed at the same center (50/65 [78.1%] in men, 23/32 [74.2%] in premenopausal women, 46/67 [68.7%] in postmenopausal women).

### Study Population

Active patient charts (N=379) of the postpolio clinic were reviewed between October 2003 and January 2004. Inclusion criteria for the study were (1) history and physical examination consistent with previous paralytic polio and (2) bone densitometry result available. Reasons for noninclusion were (1) presence of other medical disorders that can cause osteoporosis (eg, untreated thyroid disease, Paget’s disease, primary hyperparathyroidism, Cushing’s syndrome, gastrectomy, malabsorption syndrome), (2) current or previous use of medications, which can cause osteoporosis (eg, steroids, certain anticonvulsants), and (3) presence of other significant neurologic difficulties (other than paralytic polio and postpolio syndrome).

### Data Collection and Outcome Measures

The dependent variables were bone density at the femoral neck and lumbar spine (in g/cm²) and the presence or absence of osteoporosis and osteopenia at the femoral neck and lumbar spine. Osteoporosis and osteopenia were defined as per the World Health Organization (WHO) criteria (a T score at or below −2.5 for osteoporosis and a T score between −1 and −2.5 for osteopenia). Bone density of the hip was assessed arbitrarily on the left, if possible. Some patients had bone density assessments performed only at 1 site because of the presence of a total hip arthroplasty, spinal fusion, severe degenerative changes, and other causes.

Data on the following independent variables were obtained: age at evaluation, sex, BMI, age at polio, weakness at acute polio, age at menopause, history of fractures (yes, no), smoking history (yes, no), alcohol abuse history (yes, no), mobility, and current muscle strength. Weakness at acute polio was assessed on a 0 to 6 scale, with 6 being most severe (0, no weakness; 0.5, partial weakness; 1, complete paralysis for each of 4 limbs; respiratory involvement: 0, no involvement; 0.5, partial involvement; 1, use of iron lung or respirator; speech/swallowing dysfunction: 0, no involvement; 0.5, partial involvement; 1, significant difficulties). A similar measure has been found to be valid. Construct validity was evaluated by comparing weakness at acute polio in patients who were and were not hospitalized. Hospitalized patients were significantly weaker on this measure than those not hospitalized. Mobility was assessed as whether or not the patient walked daily and whether or not a wheelchair or scooter was used. If the age at menopause was unknown for a subject, it was arbitrarily set at age 50 years because we assumed that women were menopausal at age 50 or more as done in previous studies. BMI was dichotomized at 25, which is commonly used as the definition of obesity. Menopausal status was not entered into the equation for women because in many cases menopausal status was unknown and women aged 50 or more were deemed to be menopausal. Sex-specific models showed a different effect of menopausal. Sex-specific models showed a different effect of menopausal status. Thus, additional variables considered for the multivariate logistic regression model were those that could be considered confounders on a substantive basis. These included the potential predictive factors previously listed. Age was dichotomized at 50 years because we assumed that women were menopausal at age 50 or more as done in previous studies. BMI was dichotomized at 25, which is commonly used as the definition of obesity. Menopausal status was not entered into the equation for women because in many cases menopausal status was unknown and women aged 50 or more were deemed to be menopausal. Sex-specific models showed a different effect of BMI in men and women, leading us to assess an interaction between sex and BMI. This interaction was also plausible from the presence of a total hip arthroplasty, spinal fusion, severe degenerative changes, and other causes.

Data on the following independent variables were obtained: age at evaluation, sex, BMI, age at polio, weakness at acute polio, age at menopause, history of fractures (yes, no), smoking history (yes, no), alcohol abuse history (yes, no), mobility, and current muscle strength. Weakness at acute polio was assessed on a 0 to 6 scale, with 6 being most severe (0, no weakness; 0.5, partial weakness; 1, complete paralysis for each of 4 limbs; respiratory involvement: 0, no involvement; 0.5, partial involvement; 1, use of iron lung or respirator; speech/swallowing dysfunction: 0, no involvement; 0.5, partial involvement; 1, significant difficulties). A similar measure has been found to be valid. Construct validity was evaluated by comparing weakness at acute polio in patients who were and were not hospitalized. Hospitalized patients were significantly weaker on this measure than those not hospitalized. Mobility was assessed as whether or not the patient walked daily and whether or not a wheelchair or scooter was used. If the age at menopause was unknown for a subject, it was arbitrarily set at age 50 years, as done in previous studies.

Motor strength scores were calculated by using data obtained from the first postpolio clinic neurologic examination, which used a standardized form. Motor strength was evaluated by manual muscle testing by using the Medical Research Council’s (MRC) 0 to 5 scale by the same physician for all patients. Muscle strength was assessed in 7 bilateral lower-extremity muscles (hip flexors, hip extensors, hip abductors, knee extensors, knee flexors, ankle dorsiflexors, ankle plantarflexors). A sum score for each and both lower extremities was calculated for each patient by using the sum of the manual muscle testing results for hip flexors, knee extensors, and ankle dorsiflexors. A similar measure that used MRC-based physician muscle strength examination in 4 extremities has been found to have excellent interrater reliability in Guillain-Barré syndrome, a chronic neurologic disorder characterized by weakness and sensory deficits. Our sum score used the same 3 bilateral lower-extremity muscles.

### Statistical Analysis

We calculated the occurrence of osteopenia and osteoporosis at the hip and lumbar spine (as defined by the WHO criteria) in 3 patient groups (men, premenopausal women, postmenopausal women). The association between bone density measurement at the hip and muscular strength was assessed by calculation of Pearson correlation coefficients for the same 3 groups (men, premenopausal women, postmenopausal women). Confidence intervals (CIs) for the correlation coefficients were calculated by using the Fisher r-to-z transformation. For the univariate analyses of potential predictive factors in included and not included patients and in patients with and without osteoporosis at the hip, we used t tests for continuous variables and chi-square tests for dichotomous variables. When the expected number of subjects in a cell was less than 5, we used a Fisher exact test for dichotomous variables. We present P values that were not adjusted for multiple comparisons. These tables are descriptive in nature, and the P values are used to indicate which factors might be important to include in the multivariate analyses. We compared pre- and postmenopausal women with osteoporosis and osteopenia at the hip and lumbar spine with a chi-square test or Fisher exact test. A logistic regression model was estimated in the entire population of subjects to evaluate the association of osteoporosis at the hip with lower-extremity strength after adjusting for other important predictive factors. Thus, additional variables considered for the multivariate logistic regression model were those that could be considered confounders on a substantive basis. These included the potential predictive factors previously listed. Age was dichotomized at 50 years because we assumed that women were menopausal at age 50 or more as done in previous studies. BMI was dichotomized at 25, which is commonly used as the definition of obesity. Menopausal status was not entered into the equation for women because in many cases menopausal status was unknown and women aged 50 or more were deemed to be menopausal. Sex-specific models showed a different effect of BMI in men and women, leading us to assess an interaction between sex and BMI. This interaction was also plausible from the presence of a total hip arthroplasty, spinal fusion, severe degenerative changes, and other causes.

Data on the following independent variables were obtained: age at evaluation, sex, BMI, age at polio, weakness at acute polio, age at menopause, history of fractures (yes, no), smoking history (yes, no), alcohol abuse history (yes, no), mobility, and current muscle strength. Weakness at acute polio was assessed on a 0 to 6 scale, with 6 being most severe (0, no weakness; 0.5, partial weakness; 1, complete paralysis for each of 4 limbs; respiratory involvement: 0, no involvement; 0.5, partial involvement; 1, use of iron lung or respirator; speech/swallowing dysfunction: 0, no involvement; 0.5, partial involvement; 1, significant difficulties). A similar measure has been found to be valid. Construct validity was evaluated by comparing weakness at acute polio in patients who were and were not hospitalized. Hospitalized patients were significantly weaker on this measure than those not hospitalized. Mobility was assessed as whether or not the patient walked daily and whether or not a wheelchair or scooter was used. If the age at menopause was unknown for a subject, it was arbitrarily set at age 50 years, as done in previous studies.

Motor strength scores were calculated by using data obtained from the first postpolio clinic neurologic examination, which used a standardized form. Motor strength was evaluated by manual muscle testing by using the Medical Research Council’s (MRC) 0 to 5 scale by the same physician for all patients. Muscle strength was assessed in 7 bilateral lower-extremity muscles (hip flexors, hip extensors, hip abductors, knee extensors, knee flexors, ankle dorsiflexors, ankle plantarflexors). A sum score for each and both lower extremities was calculated for each patient by using the sum of the manual muscle testing results for hip flexors, knee extensors, and ankle dorsiflexors. A similar measure that used MRC-based physician muscle strength examination in 4 extremities has been found to have excellent interrater reliability in Guillain-Barré syndrome, a chronic neurologic disorder characterized by weakness and sensory deficits. Our sum score used the same 3 bilateral lower-extremity muscles.

### Results

After completion of the review, an opportunity sample of 164 patients was included in the study. A total of 215 patients were not included (159/215 [74%]) primarily because of the unavailability of bone densitometry results. Other reasons for noninclusion were the presence of other neurologic disorders (12/215 [5.6%]), other medical disorders (Cushing’s syndrome, hyperparathyroidism) (5/215 [2%]), and a diagnosis not compatible with poliomyelitis (14/215 [7%]). To evaluate the similarity of patients included in the study to the more general postpolio clinic population, patients included were compared with those not included. Men included in the...
study were significantly weaker than those not included (mean strength sum score in both legs ± standard deviation [SD], 16.8±8.3 vs 23±7.8; P.<.001). Premenopausal women included in the study were significantly older than those not included (44.5±4.9 vs 41.2±7.3y, P.<.02). They were also significantly weaker (strength sum score in both legs, 16.1±9.7 vs 21.3±9.5; P.<.02). Postmenopausal women included were also weaker at the time of the acute polio than those not included (weakness at polio score, 2.9±1.7 vs 2.2±1.5; P.<.02) and were weaker at time of the evaluation (strength sum score in both legs, 17.9±8.6 vs 21.4±8.2; P.<.01). The overall patient population included was weaker when comparing the strength sum score in both legs (P.<.05). There were no differences between men included and not included with regard to age at evaluation, weakness at polio, time since polio, and BMI. No differences were noted between premenopausal women included and not included with regard to age at polio, weakness at polio, time since acute polio, and BMI. There were no differences between postmenopausal women included and not included with regard to age at evaluation and at polio, time since acute polio, and BMI. We also compared patients included with those not included because of the unavailability of bone density results. Similar results were noted between these 2 groups to those described earlier with the exception that men included were significantly older than those not included (58±10.8y vs 53.3±11.9y, P.<.03), and premenopausal women included were not significantly weaker (strength sum score in both legs, 16.1±9.7 vs 20.2±9.8; P.<.09).

The occurrence of osteoporosis and osteopenia at the hip is as follows: 20 (32%) of 62 and 26 (40%) of 62 in men, 3 (9%) of 33 and 13 (39%) of 33 in premenopausal women, and 18 (27%) of 67 and 28 (42%) of 67 in postmenopausal women, respectively. The occurrence of osteoporosis and osteopenia at the lumbar spine is as follows: 6 (10%) of 61 and 17 (28%) of 61 in men, 2 (6%) of 32 and 7 (22%) of 32 in premenopausal women, and 7 (11%) of 65 and 20 (31%) of 65 in postmenopausal women, respectively. Osteoporosis at the hip tended to occur more frequently in postmenopausal women than premenopausal women (P.<.07), but there were no differences in the occurrence of osteoporosis at the lumbar spine and in the occurrence of osteopenia in premenopausal women compared with postmenopausal women.

In univariate analyses (table 1), in men, moderate correlations were found between hip bone density and hip flexor strength (Pearson correlation coefficient, r=27; 95% CI, 0.23–0.49) and between hip bone density and strength sum score (hip flexors, knee extensors, ankle dorsiflexors) in the same extremity in which the bone density was performed (Pearson r=0.32; 95% CI, 0.07–0.53). In premenopausal women, hip bone density and strength sum score in the same lower extremity correlated moderately (r=0.35; 95% CI, 0.17–0.50), and, in postmenopausal women, moderate correlations were found between hip bone density and same hip flexor strength (r=0.3; 95% CI, 0.17–0.50), strength sum score in the same lower extremity (r=0.25; 95% CI, 0.12–0.46), and strength sum score in both lower extremities (r=0.27; 95% CI, 0.03–0.48). A comparison of potential predictive factors in patients with and without osteoporosis at the hip is presented in tables 2 and 3. Premenopausal women with osteoporosis were younger at acute polio, and postmenopausal women with osteoporosis were weaker at acute polio, but there were no other differences between patients with and without osteoporosis with regard to age and weakness at acute polio, time since acute polio, current age, and BMI. There were also no differences between patients with and without osteoporosis in terms of mobility status and history of smoking and alcohol abuse. Because there were only 3 premenopausal women with osteoporosis, we also compared potential predictive factors for osteoporosis in premenopausal women with and without osteopenia at the hip. These results were similar to those for osteoporosis with the exception that premenopausal women with osteopenia did not differ from those without osteopenia with regard to age at acute polio.

In a multivariate model (table 4), osteoporosis at the hip was significantly associated with strength sum score in the same lower extremity (odds ratio [OR], .42 per 5-point increase in strength score; 95% CI, .28–.64; P<.001) after adjusting for time since acute polio (OR=1.04; 95% CI, 1.00–1.09; P=.03), age at evaluation less than 50 years (OR=0.47; 95% CI, 0.14–1.50; P=.20), male sex (OR=0.87; 95% CI, 0.28–2.62; P=.80), and interaction between weight and sex (P=.08).
found a lower frequency of osteoporosis at the lumbar spine in densitometry in a postpolio clinic population. However, we pausal and postmenopausal women who are referred for bone osteopenia at the hip were common in men and in premeno-

### Table 2: Comparison of Potential Predictive Factors in Patients With and Without Osteoporosis at the Hip: Categorical Variables

<table>
<thead>
<tr>
<th>Potential Predictive Factors</th>
<th>Osteoporosis</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes (%)</td>
<td>No (%)</td>
<td>P</td>
<td></td>
</tr>
<tr>
<td><strong>Men</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mobility status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walking daily</td>
<td>19/20 (95.0)</td>
<td>35/42 (83.3)</td>
<td>0.26</td>
<td></td>
</tr>
<tr>
<td>Use of wheelchair or scooter</td>
<td>5/20 (25.0)</td>
<td>10/42 (23.8)</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>History of smoking</td>
<td>4/20 (20.0)</td>
<td>6/42 (14.3)</td>
<td>0.71</td>
<td></td>
</tr>
<tr>
<td>History of alcohol abuse</td>
<td>1/20 (5.0)</td>
<td>1/42 (2.4)</td>
<td>0.54</td>
<td></td>
</tr>
<tr>
<td><strong>Premenopausal women</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mobility status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walking daily</td>
<td>1/3 (33.3)</td>
<td>22/30 (73.3)</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td>Use of wheelchair or scooter</td>
<td>2/3 (66.7)</td>
<td>9/30 (30.0)</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>History of smoking</td>
<td>1/3 (33.3)</td>
<td>8/30 (26.7)</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>History of alcohol abuse</td>
<td>0/3 (0.0)</td>
<td>1/30 (3.3)</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td><strong>Postmenopausal women</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mobility status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walking daily</td>
<td>12/18 (66.7)</td>
<td>44/49 (89.8)</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>Use of wheelchair or scooter</td>
<td>6/18 (33.3)</td>
<td>13/49 (26.5)</td>
<td>0.76</td>
<td></td>
</tr>
<tr>
<td>History of smoking</td>
<td>4/18 (22.2)</td>
<td>6/49 (12.2)</td>
<td>0.44</td>
<td></td>
</tr>
<tr>
<td>History of alcohol abuse</td>
<td>0/18 (0.0)</td>
<td>1/49 (2.0)</td>
<td>1.00</td>
<td></td>
</tr>
</tbody>
</table>

Being overweight appeared to be harmful for men (OR = 2.4; 95% CI, 0.71–8.00) and protective for women (OR = 0.57; 95% CI, 0.19–1.71).

### DISCUSSION

In this exploratory study, we found that osteoporosis and osteopenia at the hip were common in men and in premenopausal and postmenopausal women who are referred for bone densitometry in a postpolio clinic population. However, we found a lower frequency of osteoporosis at the lumbar spine in our population. A significant correlation was found between bone mass and muscle strength (strength sum score in the same lower extremity in which bone densitometry was performed) in both our univariate and multivariate analyses. We did not find other significant associations between bone mass and other known risk factors such as smoking, alcohol consumption, age, and sex.

To our knowledge, this is the most comprehensive study of osteoporosis in postpolio patients to date. Osteoporosis is likely underdiagnosed in this population and in other neuromuscular disorders, especially in men and premenopausal women. Osteoporosis is mostly associated with older patients, primarily postmenopausal women, leaving this group of high-risk patients without appropriate treatment. Our findings of a high frequency of osteoporosis in a postpolio clinic population and a significant correlation of bone mass with lower-extremity muscular strength should alert clinicians and prompt them to systematically refer these patients for bone densitometry.

We found that osteoporosis at the hip was common in a postpolio clinic population in men (32%) and in premenopausal (12.5%) and postmenopausal women (27%). These numbers are higher than those for the general population for men at the hip in Canada (4.8%) and for postmenopausal women at the hip in Canada (7.9%). We found that the frequency of osteoporosis is very high in a postpolio clinic patient population in 3 different groups (men, 45%; premenopausal women, 37.5%; postmenopausal women, 34%). Therefore, these patients are at high risk for developing osteoporosis in the future. It has been shown that the fracture risk increases 1.5- to 3-fold for each SD decrease in BMD and that, at any given age, the lifetime risk of a proximal femoral fracture rises as bone density diminishes.

Surprisingly, the T-score values obtained at the lumbar spine in our population did not show the same prevalence for osteoporosis as at the hip. One explanation for this finding would be that most of our patients likely have degenerative disorders of the spine, including scoliosis in some. These musculoskeletal abnormalities can likely falsely elevate the T score at the lumbar spine, as described in previous studies. Another explanation could be that in this patient population osteoporosis may be a regional phenomenon; the primary predictive factor for osteoporosis at the hip was weakness in the same lower extremity. Therefore, those patients with lower-extremity weakness are more susceptible to regional osteoporosis affecting mostly the hip rather than diffuse osteoporosis affecting both the lumbar spine and the hip. Similar findings have been reported in the spinal cord injury population. These pa-

### Table 3: Comparison of Potential Predictive Factors in Patients With and Without Osteoporosis at the Hip: Continuous Variables

<table>
<thead>
<tr>
<th>Potential Predictive Factors</th>
<th>Osteoporosis</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
<td>P</td>
</tr>
<tr>
<td><strong>Men</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age at acute polio (y)</td>
<td>5.7 ± 5.5</td>
<td>7.5 ± 9.8</td>
<td>.34</td>
</tr>
<tr>
<td>Weakness at acute polio (0–6 scale)</td>
<td>1.9 ± 1.3</td>
<td>2.5 ± 1.5</td>
<td>.15</td>
</tr>
<tr>
<td>Time since acute polio (y)</td>
<td>55.3 ± 12.1</td>
<td>49.2 ± 11.8</td>
<td>.07</td>
</tr>
<tr>
<td>Current age (y)</td>
<td>61.0 ± 11.3</td>
<td>56.8 ± 10.7</td>
<td>.16</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>25.6 ± 4.1</td>
<td>26.2 ± 5.4</td>
<td>.64</td>
</tr>
<tr>
<td><strong>Premenopausal women</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age at acute polio (y)</td>
<td>2.3 ± 0.6</td>
<td>5.7 ± 8.1</td>
<td>.03</td>
</tr>
<tr>
<td>Weakness at acute polio (0–6 scale)</td>
<td>2.6 ± 2.1</td>
<td>2.2 ± 1.5</td>
<td>.67</td>
</tr>
<tr>
<td>Time since acute polio (y)</td>
<td>41.7 ± 2.4</td>
<td>38.9 ± 9.6</td>
<td>.62</td>
</tr>
<tr>
<td>Current age (y)</td>
<td>44.0 ± 2.2</td>
<td>44.6 ± 4.8</td>
<td>.85</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>25.6 ± 4.6</td>
<td>25.8 ± 3.9</td>
<td>.96</td>
</tr>
<tr>
<td><strong>Postmenopausal women</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age at acute polio (y)</td>
<td>9.1 ± 9.3</td>
<td>5.7 ± 4.9</td>
<td>.15</td>
</tr>
<tr>
<td>Weakness at acute polio (0–6 scale)</td>
<td>3.7 ± 1.7</td>
<td>2.6 ± 1.6</td>
<td>.01</td>
</tr>
<tr>
<td>Time since acute polio (y)</td>
<td>56.3 ± 13.3</td>
<td>56.2 ± 8.9</td>
<td>.97</td>
</tr>
<tr>
<td>Current age (y)</td>
<td>58.9 ± 11.4</td>
<td>62.0 ± 8.5</td>
<td>.17</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>26.3 ± 6.2</td>
<td>26.6 ± 5.3</td>
<td>.86</td>
</tr>
</tbody>
</table>

NOTE. Values are n or mean ± SD.

### Table 4: Adjusted (multivariate model) ORs for Osteoporosis at the Hip

<table>
<thead>
<tr>
<th>Variable</th>
<th>OR</th>
<th>95% CI</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strength sum score in same lower extremity (per 5-point increase in strength score)</td>
<td>0.42</td>
<td>0.28–0.64</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Age at evaluation (&lt;50y)</td>
<td>0.47</td>
<td>0.14–1.5</td>
<td>.20</td>
</tr>
<tr>
<td>Overweight (BMI &gt;25 kg/m²)</td>
<td>0.71–8.0</td>
<td>.08*</td>
<td></td>
</tr>
<tr>
<td>For men</td>
<td>2.37</td>
<td>1.03–1.089</td>
<td>.03</td>
</tr>
<tr>
<td>Male sex</td>
<td>0.87</td>
<td>0.28–2.62</td>
<td>.80</td>
</tr>
</tbody>
</table>

NOTE. Model is based on 164 patients included in the study, with and without osteoporosis at the hip.

*P value for interaction.
tients also have a higher prevalence of osteoporosis at the hip compared with the vertebral spine likely related to immobilization.\textsuperscript{19,20} Nonetheless, osteoporosis, even if localized primarily to the hip, remains an important risk factor for hip fractures.\textsuperscript{6,16,21} In fact, hip fracture is most closely linked to BMD compared with other types of fractures.\textsuperscript{2,16,21}

Moreover, in this exploratory study, in our multivariate model, we found a significant association between osteoporosis at the hip and strength sum score in same extremity with an OR of .42 for each 5-point increase in the strength sum score, after adjusting for several other risk factors (time since acute polio, age at evaluation, male sex). This indicates that with each 5-point increase in the strength sum score (ie, greater strength), postpolio clinic patients were less likely to have osteoporosis (by a factor on average of .42). In addition, we found that an increased time since acute polio (with resultant muscular weakness) was a significant predictor of osteoporosis, stressing once again the protective aspect of muscle strength. Our study reflects the important role that muscle strength plays in protecting bone from osteoporosis.\textsuperscript{4,5,18} In fact, it has been shown that in the general population being physically active reduces the risk of later hip fracture by up to 50%.\textsuperscript{5,8,21,22}

An unexpected finding in our study was that being overweight appeared to be harmful for men tending to increase their risk to have osteoporosis, whereas for women being overweight tended to be protective. Several previous studies have shown that a low body weight is negatively correlated with peak bone mass.\textsuperscript{8} In fact, the more weight a woman gains since age 25, the lower her risk of hip fracture.\textsuperscript{5} In addition, higher adiposity is protective against the risk of both hip and vertebral fractures in women.\textsuperscript{8,22} However, we do not believe that previous studies have shown a difference in the association between osteoporosis and BMI in men and women.

Study Limitations

Our study had several potential limitations. The patients included in our study were from a university-affiliated hospital postpolio clinic and those that had undergone a bone densitometry as requested by the clinician physician. They were likely more severely affected by previous paralytic polio and more likely at greater risk for osteoporosis. Hence, our estimate on the prevalence of osteoporosis may be overestimated compared with the general postpolio population. The patients included in our study were in fact weaker than those not included. Nevertheless, the bone densitometry assessments performed were not always on the weaker side (the left side is usually done by convention). In fact, 19.3\% were performed on the weaker side in men, 18\% in premenopausal women, and 12\% in postmenopausal women. This could have underestimated the prevalence of osteoporosis with normal bone densitometry results when not performed on the weaker side. Our study did not find a significant correlation between bone density and alcohol or tobacco use, which are 2 recognized risk factors.\textsuperscript{2,5,8,21} This can be explained by an insufficient number of patients who used alcohol or tobacco and by the fact that the exact amount of cigarettes and alcohol consumption was not recorded. This study had potential bias because some potential predictive factors were ascertained by self-report and retrospectively. The strengths of our study are the relatively large study population, study comprehensiveness, and consistency in data acquisition and ascertainment of bone density (most patients evaluated at the same radiologic department and all patients evaluated by the same physician by using a standardized form).

CONCLUSIONS

In this retrospective, cross-sectional study, we found that osteoporosis and osteopenia at the hip occur commonly in postpolio clinic patients referred for bone densitometry in men, premenopausal women, and postmenopausal women compared with the general population. However, because patients included in the study were weaker than patients not included, our estimates of the occurrence of osteoporosis in a postpolio clinic population are likely elevated. Hip bone density was associated with muscle strength in the same lower extremity in which the examination was performed in men, premenopausal women, and postmenopausal women attending a postpolio clinic in both unadjusted and adjusted multivariate analyses. The diagnosis can therefore be missed if the bone densitometry is not performed on the same side as the weaker lower extremity. Based on these results, we recommend that all postpolio patients be evaluated for osteoporosis at both hips (or less preferably at the hip of the weaker lower extremity) and at the lumbar spine. It is possible that with treatment BMD will improve and fracture risk will decline. We recommend further research to confirm the results of our exploratory study and to evaluate the effects of treatment on osteoporosis and fracture rate in the postpolio and other neuromuscular disease patient populations.

References

Lowering of Sensory, Motor, and Pain-Tolerance Thresholds With Burst Duration Using Kilohertz-Frequency Alternating Current Electric Stimulation

Alex R. Ward, PhD, Stacey Lucas-Toumbourou, MPH


Objective: To determine the optimum burst duration for discrimination between sensory, motor, and pain-tolerance thresholds using 50-Hz bursts of kilohertz-frequency sinusoidal alternating current (AC) applied transcutaneously to human subjects.

Design: A repeated-measures randomized controlled trial.

Setting: A research laboratory.

Participants: Twenty-six healthy young adults.

Interventions: Bursts of AC electric stimulation at frequencies of 1 and 4kHz. Burst durations ranged from 250μs (for 1 cycle of 4kHz AC, ie, a single biphasic pulse) to 20ms (continuous AC).

Main Outcome Measures: We measured sensory, motor, and pain-tolerance thresholds at frequencies of 1 and 4kHz.

Results: We found that threshold voltages decreased to a minimum with increasing burst duration. The minimum threshold identified the “utilization time” over which summation of subthreshold stimuli occurs. Above the utilization time, thresholds increased. Estimated utilization times differed for sensory (≅7ms), motor (≅10ms), and pain-tolerance (≅20ms). As a consequence, relative thresholds varied with burst duration. A maximum separation between sensory, motor, and pain-tolerance thresholds was found to occur with bursts in the range 1 to 4ms.

Conclusions: Short-duration kilohertz-frequency AC bursts might have a more useful role in rehabilitation than either pulsed current or the long duration bursts that characterize Russian and interferential currents. Further clinical studies are needed.

Key Words: Electric stimulation; Pain threshold; Rehabilitation.

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

TRANSCUTANEOUS ELECTRIC stimulation is commonly used for pain control and for benefits resulting from muscle contraction such as muscle re-education (modulating sensory-motor neural connectivity), minimizing muscle atrophy, and strengthening. Figure 1 illustrates the difference between the 2 stimuli types commonly used in clinical practice: short-duration pulsed current and kilohertz-frequency alternating current (AC). Pulsed current stimulators typically deliver pulse durations in the range 20 to 400μs and a pulse frequency between 1 and 120Hz. They may be biphasic (fig 1) or monophasic. AC stimulation typically uses a symmetrical biphasic waveform (sinusoidal or rectangular in shape) with frequencies between 1 and 10kHz, applied in bursts with a burst frequency between 1 and 120Hz and a burst duty cycle of 10% or more.²,³

The widespread use of kilohertz-frequency AC in rehabilitation, particularly in the form of so-called interferential current and Russian current, seems to be due to unsubstantiated clinical belief that AC stimulation is more comfortable and more effective than pulsed current.¹,⁴ These clinical presumptions have, until recently,³,⁸ not been questioned, but there is some historical evidence.⁹,¹⁵

Kilohertz-Frequency AC Stimulation and Summation

Unlike the pulsed current stimulation used clinically, with kilohertz-frequency AC stimulation there is the possibility of summation of subthreshold depolarizations: a phenomenon first described by Gildemeister.₁²,₁₃ Summation occurs when a burst of AC of sufficiently high frequency is used as the stimulus. During each burst, the nerve fiber membrane is pushed closer to threshold with each successive pulse in the wave-train because the nerve does not have time to recover between pulses. Eventually threshold is reached and an action potential is produced. This means that short-duration bursts will allow for little summation, so the thresholds will be high. Longer duration bursts allow greater summation so the threshold will be lower. Based on his observations of somatosensory response, Gildemeister concluded that there was a limit to the time over which pulses could summate: a plateau is reached where further increases in burst duration do not result in any further decrease in threshold voltage. He reported this as the “nutzeit” or “utilization time” over which summation can occur.

Schwarz and Ehrig used frog and rat nerve trunk preparations to measure compound action potentials in response to single bursts of kilohertz-frequency AC with 1 to 30 cycles per burst. AC frequencies used were in the range 770Hz to 40kHz. They found that threshold stimulus voltages decreased as the number of cycles per burst increased. They observed a nuttzeit, a plateau in their graphs of threshold versus burst duration, that was the same (≅2–4ms) for all the AC frequencies used.

Later work by Schwarz and Volkmer used single fiber preparations to directly measure the change in membrane potential at a node of Ranvier when successive pulses of a 2.4kHz-frequency AC stimulus were applied. They found that...
Discomfort and Kilohertz-Frequency AC Stimulation

A recent study, which assessed discomfort over the frequency range of 0.5 to 20kHz, found that discomfort decreased as frequency increased from 0.5 to 4 to 5kHz then increased as frequency increased to 20kHz. The observation that maximum discomfort occurs at a frequency of 4 to 5kHz helps to justify that AC stimulation is advantageous in that it evokes least discomfort in this frequency range.

Burst duration is also an important factor determining discomfort. Ozcan et al\(^\text{6,8}\) reported that continuous 4kHz AC and “true” interferential current (with a beat frequency of 50Hz) caused more discomfort and elicited less quadriceps muscle torque than 10ms, 50-Hz bursts of 4-kHz AC (as used in premodulated interferential stimulation). Ward et al\(^\text{8}\) examined a range of AC frequencies (0.5–20kHz) at different burst durations to stimulate the wrist extensors and found that a burst duration of about 4ms resulted in least discomfort and that maximum muscle torque is elicited using bursts of about 2ms. Burst duration appeared to be the important factor and not the number of cycles per burst.

The Present Study

Although it has been found\(^\text{6,8}\) that short burst durations (1–4ms) are associated with greatest torque production and least discomfort, no convincing explanation has yet been advanced regarding this phenomenon. A likely explanation is that the different nerve fiber types associated with sensory stimulation, motor activation, and discomfort respond differently to AC burst stimulation because of their different electrophysiologic properties. Different nerve fiber types have different refractory periods, different periods of hypo- and hyperexcitability, and different recovery times.\(^\text{18-20}\) This implies that sensory, motor, and pain fibers will respond differently to stimuli with varying burst duration. The summation rate and amount of summation would be expected to vary between fiber types and, at burst frequencies around 50Hz (as commonly used clinically for motor stimulation), the time between bursts (the interburst interval) could be within the refractory period and so affect the measured threshold. Thus it might well be that relative discomfort and stimulus-elicited torque depend on burst duration because of these factors.

The present study was intended to provide some insight into the importance of the factors described above. We examined whether sensory, motor, and pain responses are affected to the same extent if stimulus characteristics, specifically the burst duration and AC frequency, are changed.

METHODS

The 26 participants in this study were volunteers who responded to notice-board advertisements and met the inclusion criteria. That is, they did not have a pacemaker (or indwelling stimulator) or any breaks in the skin under the area where the electrodes were to be placed, and had no known neurologic or musculoskeletal pathologies affecting the upper limb to be tested. The group of participants (11 men, 15 women) included academic staff members and undergraduate students of the university (age range, 19–56y; mean, 23y). Approval was obtained from the Ethics Committee of the Faculty of Health Sciences of La Trobe University prior to the commencement of this study.

After we explained the procedure to each participant, and obtained informed consent, the skin of the dorsal surface of the left forearm was cleaned with an alcohol swab, and conductive rubber electrodes, measuring 37×45mm, were attached using conductive, adhesive skin mounts.\(^\text{7}\) The electrodes were positioned to stimulate the wrist extensors: 1cm apart on a line from the head of the radius to the distal radio-ulnar joint with the proximal electrode 1cm distal to the head of the radius and the distal electrode 5cm distal to the proximal electrode along this line.

The stimulator was a purpose-built device designed to produce constant voltage stimuli consisting of bursts of sinewaves with user selection of the sinewave frequency, burst frequency, and number of sinewaves per burst. Calibration accuracy of both the sinewave frequency and burst frequency was ±2% and there was no discernible variation in either parameter over the time course of the testing sessions.

Participants experienced stimuli at 2 sinewave frequencies (1kHz, 4kHz), applied in bursts at a burst frequency of 50Hz. Stimulus waveforms had a predetermined number of sinewave cycles per burst and therefore a predetermined burst duration. At 1kHz, burst durations of about 4 ms were associated with maximum force production and least discomfort, although there was considerable variation in the actual burst duration. Burst duration appeared to be the important factor and not the number of cycles per burst.

with each successive pulse, the measured membrane potential changed until, after several pulses, threshold was reached and the nerve fiber fired. This was the first reported direct observation of summation of subthreshold depolarizations.

This raises the question of whether any different physiologic effects of pulsed current and bursts of kilohertz-frequency AC (eg, relative discomfort, force production) are explicable in terms of summation. Summation occurs when successive pulses are applied before the nerve-fiber membrane can recover. With kilohertz-frequency AC there are multiple pulses, close together, in a burst so pulses can summate. The separation between pulses is hundreds of microseconds. With pulsed current used clinically, the separation between pulses is much larger (almost 10ms at a pulse frequency of 100Hz and larger at lower frequencies). Even with burst-mode transcutaneous electric nerve stimulation, the pulse frequencies in the burst are normally 150Hz or less, so that the separation between pulses is about 7ms or more, too long to allow appreciable summation of the kind described here. Thus kilohertz-frequency AC (including Russian and interferential currents) would appear to allow for summation, whereas pulsed current (as used clinically) would not.

There is a limit to summation, specified by the utilization time, which is the time over which summation can occur. When pulses are applied in bursts, a longer burst duration (larger number of pulses) results in lower nerve-fiber thresholds, that is, a greater summation effect, provided the burst duration is less than the utilization time. When the burst duration exceeds the utilization time, no further lowering of thresholds occurs.

Discomfort and Kilohertz-Frequency AC Stimulation

A recent study, which assessed discomfort over the frequency range of 0.5 to 20kHz, found that discomfort decreased as frequency increased from 0.5 to 4 to 5kHz then increased as frequency increased to 20kHz. The observation that maximum discomfort occurs at a frequency of 4 to 5kHz helps to justify that AC stimulation is advantageous in that it evokes least discomfort in this frequency range.

Burst duration is also an important factor determining discomfort. Ozcan et al\(^\text{6,8}\) reported that continuous 4kHz AC and “true” interferential current (with a beat frequency of 50Hz) caused more discomfort and elicited less quadriceps muscle torque than 10ms, 50-Hz bursts of 4-kHz AC (as used in premodulated interferential stimulation). Ward et al\(^\text{8}\) examined a range of AC frequencies (0.5–20kHz) at different burst durations to stimulate the wrist extensors and found that a burst duration of about 4ms resulted in least discomfort and that maximum muscle torque is elicited using bursts of about 2ms. Burst duration appeared to be the important factor and not the number of cycles per burst.
used (20 cycles meaning continuous stimulation, ie, no inter-burst interval). At 4kHz, 1, 2, 4, 8, 12, 16, 24, 32, 40, 48, 60, and 80 cycles per burst were used (80 cycles being continuous stimulation and so no interburst interval). The 21 different combinations of AC frequency and cycles per burst were applied in a randomized order for each participant. Randomization was achieved by creating a column of random numbers in an Excel spreadsheet, which listed the combinations of frequency and number of sinewaves, then sorting the list by the random number column.

Once the randomized order was determined, we used it to determine sensory thresholds. Participants were asked to manually increase the stimulus intensity until the first perception of cutaneous sensation was reached. The experimenter recorded this threshold for 3 measurements at each particular frequency and number of cycles per burst, for each of the 21 combinations of sinewave frequency and burst duration. Once all sensory measurements had been obtained, motor thresholds were determined in the same sequence as used for sensory thresholds. The motor threshold was taken as the first overt indication of muscle activity: visible contraction under the electrodes or wrist or finger extension noted by the experimenters. Finally, pain-tolerance thresholds were measured. Participants increased the intensity to a value where they felt that any further increase would be too painful. Although pain tolerance is a subjective measure, the repeated-measures design allowed between-subject variation to be factored out so that it was not a confounding variable. All measurements were taken in a single experimental session of approximately 1 hour.

Prior to measurements at each threshold, participants experimented with the intensity control to familiarize themselves with the stimulus and establish the criterion which would be used to determine each threshold.

RESULTS

We averaged the 3 measurements of each threshold (at each particular frequency and number of cycles per burst) for each participant, then averaged over the group of participants. Figure 2 shows the group-averaged variation in sensory, motor, and pain-tolerance thresholds with burst duration (in milliseconds) at 1 and 4kHz. The error bars shown are standard deviations. A surprising feature is that the graphs appear to reach a minimum threshold rather than simply decreasing to a plateau. This is most clearly evident in figure 2A where a minimum sensory threshold is reached at a burst duration of 5 to 7ms for both 1 and 4kHz AC. Motor thresholds seem to be minimal at a little over 10ms (fig 2B) and pain-tolerance thresholds, more difficult to discern, possibly about 20ms (fig 2C). Again the variation is similar at 1 and 4kHz.

The error bars shown in figure 2 are large and this calls into question whether the apparent variations in threshold with burst duration are real or spurious. To test whether the variations were significant, we performed repeated-measures analyses of variance (ANOVA) separately for each threshold and kilohertz frequency. All 6 ANOVA found a significant between-subject effect (table 1). F values were in the range 149 to 302 and the corresponding P values were all .000. The variation with burst-duration was also significant: F values were in the range 14 to 66 and the corresponding P values were all .000.

The higher F values for the between-subject variance indicate that the large error bars in figure 2 are more due to between-subject variance than between burst-duration variance. With the between-subject variance subtracted, the burst duration variation apparent in figure 2 is found to be significant (P = .000 for each threshold and kilohertz frequency) (see table 1).

If the minimum thresholds for sensory, motor, and pain responses occur at different burst durations (as suggested by fig 2), relative thresholds, in particular pain divided by sensory and pain divided by motor, would be expected to vary and there will be an optimal burst duration where maximum discrimination between thresholds is achieved. Because each threshold...
appears to have a different minimum, the optimal burst duration could depend on the thresholds being compared. To explore and test these predictions, we calculated relative thresholds (pain threshold/sensory threshold and pain threshold/motor threshold) for each burst duration at each kilohertz frequency (fig 3).

The graph of pain threshold/sensory threshold (see fig 3A) shows little difference between the 1- and 4-kHz results. A distinct maximum is seen in the 4-kHz results, at a burst duration of about 2ms. At 1kHz, it is difficult to establish a maximum as the shortest burst duration possible at this frequency is 1ms but the consistency with the 4-kHz results supports the notion of bursts of about 2ms being optimal for discrimination between pain and sensory responses at both kilohertz frequencies. The graphs of pain threshold/motor threshold (see fig 3B) also appear to show a maximum at about 2 to 3ms but the apparent variation is less convincing. Interpretation of the 1-kHz results is difficult because, as noted earlier, the shortest burst duration possible at this frequency is 1ms. It should also be noted that more error is introduced when relative thresholds are calculated, because the errors compound. The data in figure 3 thus have about double the scatter of those in figure 2. This is reflected in the size of the error bars. A conservative and rational conclusion is therefore that the 1- and 4-kHz results probably do not differ and that best discrimination between each threshold appears to occur at about 1- to 4-ms burst duration.

To test whether the apparent variation of relative threshold with burst duration (see fig 3) was significant, we conducted repeated-measures ANOVAs for each relative threshold and each kilohertz frequency (4 ANOVAs). The results are summarized in table 2. All 4 ANOVAs found a significant between-subject effect. F values are in the range 45 to 122 and the corresponding P values are all .000. The burst-duration effect was significant at 4kHz and also at 1kHz for the pain/sensory thresholds. F values were in the range 2.6 to 8.9 and the corresponding P values are all less than .003. The 1-kHz pain/motor relative threshold did not show a statistically significant variation (F = 1.6) but the P value (.113) suggests that if the null hypothesis is accepted (ie, no variation with burst duration), this conclusion has an 89% probability of being wrong, that is, a type II error: accepting the null hypothesis when in fact there is actually a difference.

Again, the higher F values for the between-subject variance indicate that the large error bars in figure 3, as in figure 2, are more due to between-subject variance than between burst-duration variance.

The repeated-measures ANOVAs that found a statistically significant variation in relative threshold with burst duration do not indicate which particular burst duration or range of burst durations is “best.” Post hoc tests are needed. With a repeated-measures design and multiple burst durations, it is not appropriate to conduct post hoc comparisons of every burst duration with all of the others. Because 10 burst durations were used at 1kHz and 12 at 4kHz, there are 45 possible pairwise compa-
isons at 1kHz and 66 at 4kHz. The Bonferroni adjustment factor to the acceptable \( P \) value would be large, that is, rather than accepting significance at the .05 level, the level would have to be .05/45 = .001 for the 1-kHz results and .05/66 = .0008 for the 4-kHz results. This would increase the risk of a type II statistical error to an unacceptably high level.

Inspection of figure 3 indicates that a burst duration in the range 1 to 4ms could possibly be optimal. Accordingly data for burst durations in the range 1 to 4ms was averaged and compared with averaged data for all longer burst durations (6–20ms) and with averaged data for all shorter burst durations (.25ms and .50ms at 4 kHz). This reduced the number of post hoc comparisons to 1 (at 1kHz) or 2 (at 4kHz), so the confounding effect of a large Bonferroni adjustment was minimized. Paired \( t \) tests showed that at both 1 and 4kHz the pain/motor and pain/sensory thresholds were significantly higher for 1- to 4ms bursts as compared with longer burst durations \( (P = .000) \). Comparison of the 4-kHz results at 1 to 4ms with those at .25 to .50ms found no significant difference \( (P = .109) \). The small \( P \) value, however, indicates that the chance of a type II error is high if one accepts the null hypothesis. The low \( P \) value suggests that there is a small difference that would only be convincingly shown with larger numbers of participants.

**DISCUSSION**

Our finding that thresholds decrease with increasing burst duration is unsurprising: the phenomenon was first reported by Gildean9 between 1930 and it became known as the Gildemeister effect, whereby successive subthreshold stimuli summate and threshold is reached by summation of subthreshold stimuli. Summation of subthreshold stimuli has been shown more recently using noninvasive12,13 and invasive16,17 techniques. Our data reaffirm this established electrophysiologic phenomenon as the peak stimulus intensity decreased with increasing number of AC cycles (and hence burst duration) until a minimum, the utilization time (or Nuttzeit), was reached. The idea is that when the nerve fiber membrane undergoes subthreshold depolarization, the resulting ion concentration changes take time to recover. Extra stimulus pulses occurring before recovery can therefore summate, pushing ion concentrations further from their resting values until threshold is reached and an action potential is produced. So summation can occur, but only if additional stimulus pulses are delivered within the recovery time (the Nuttzeit).

We found that sensory, motor, and pain-tolerance thresholds appeared to vary differently with burst duration (see fig 2). Our findings indicate that the utilization time is different for sensory, motor, and pain-tolerance thresholds. This suggests that the utilization time is shortest for large diameter sensory fibers, longer for equal or smaller diameter motor fibers, and longest for pain fiber stimulation. A difference in utilization time means that pain/sensory and pain/motor threshold ratios will vary with burst duration. This suggests the possibility that there will be optimal burst durations where maximum sensory or motor responses can be elicited with minimal risk of pain. A high pain/sensory or pain/motor threshold indicates better separation between the thresholds and so a greater ability to evoke a sensory or motor response without producing pain. Figure 3 indicates that for both pain/sensory and pain/motor discrimination, a burst duration in the range of 1 to 4ms is desirable.

Three major findings of the present study are thus that: (1) summation of subthreshold stimuli results in bursts of AC producing lower thresholds than single, biphasic pulses; (2) there is a minimum in the threshold/burst-duration graph that is different for sensory, motor, and pain-tolerance stimulation (see fig 2); and (3) as a consequence, there is best discrimination between pain versus motor and pain versus sensory stimulation at particular burst durations, identified in figure 3 as in the range of 1 to 4ms.

The optimal discrimination at low (1–4ms) burst durations (see fig 3) suggests that greater sensory and motor stimulation before evoking a pain response is possible at shorter burst durations. This helps to explain earlier findings8 that greatest muscle torque is elicited, and that least discomfort also results, by using short duration bursts. Ward et al used 50-Hz bursts of AC to compare a range of AC frequencies and duty cycles (burst durations) and found that 1-kHz AC with a low burst duration (1–2ms) was optimal for torque production whereas a higher frequency and burst duration (4-kHz AC, 2–4ms bursts) was optimal for minimizing discomfort. These optimal burst durations (1–4ms) are much shorter than the 10ms typically used clinically for sensory and motor stimulation8 and adapting them may lead to more comfortable perception of the AC stimulation.

The present study also helps to explain why stimulation using bursts of kilohertz-frequency AC is more comfortable than stimulation using pulsed current.1,3,7,8 The shortest burst duration with kilohertz-frequency AC is that of 1 AC cycle, which is actually a biphasic pulsed current stimulus. The 4-kHz results in figure 3 suggest that when single biphasic pulses (minimum burst duration) are used, the relative thresholds are less than with 1 to 4ms bursts. A lesser discrimination between sensory, motor, and pain thresholds with single pulses is consistent with the observation that stimulation using pulsed current is more uncomfortable than when bursts of kilohertz-frequency AC are used.1,3,7,8

A clinical implication is that for sensory stimulation there is least likelihood of evoking pain if short-duration bursts are used. Further, for motor stimulation with the least discomfort, short-duration bursts will be more effective than either single-cycle AC or long-duration bursts.

### Table 2: Results of the Repeated-Measures ANOVAs for Relative Thresholds (pain/sensory and pain/motor) With Burst Duration

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Relative Threshold</th>
<th>Factor</th>
<th>( F )</th>
<th>( P )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1kHz</td>
<td>Pain/sensory</td>
<td>Between subject</td>
<td>122</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Burst duration</td>
<td>8.9</td>
<td>.000</td>
</tr>
<tr>
<td>4kHz</td>
<td>Pain/sensory</td>
<td>Between subject</td>
<td>66.0</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Burst duration</td>
<td>4.4</td>
<td>.000</td>
</tr>
<tr>
<td>1kHz</td>
<td>Pain/motor</td>
<td>Between subject</td>
<td>44.7</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Burst duration</td>
<td>1.8</td>
<td>.113</td>
</tr>
<tr>
<td>4kHz</td>
<td>Pain/motor</td>
<td>Between subject</td>
<td>75.6</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Burst duration</td>
<td>2.6</td>
<td>.003</td>
</tr>
</tbody>
</table>
CONCLUSIONS

The mechanism associated with peripheral nerve excitation using bursts of kilohertz-frequency AC is most likely a summation effect with a distinct time frame (the utilization time) over which summation is possible. The present study also found that utilization times differed for sensory (≈5–7ms), motor (≥10ms), and pain-tolerance (≥20ms) thresholds, indicating that there is a specific burst duration where separation between sensory, motor, and pain thresholds is greatest.

A comparison of relative thresholds identified the range of 1 to 4ms as best for sensory, motor, and pain-threshold discrimination.

References
5. Laufer Y, Ries JD, Leninger PM, Alon G. Quadriceps femoris muscle torque and fatigue generated by neuromuscular electrical stimulation with three different waveforms. Phys Ther 2001;81:1307-16.

Suppliers
a. American Imex, 16520 Aston St, Irvine, CA 92714.
b. Microsoft Corp, One Microsoft Wy, Redmond, WA 98052.
Pain Perception After Running a 100-Mile Ultramarathon

Martin D. Hoffman, MD, Jean Lee, MS, MD, Holly Zhao, MD, PhD, Alex Tsodikov, PhD


Objective: To determine if pain perception is affected by an extreme bout of exercise that causes ongoing exercise-related pain.

Design: Repeated-measures design.

Setting: Pre-race registration area and finish area of an endurance race.

Participants: Twenty-one competitors in the 2005 Western States 100 Mile Endurance Run and 11 control subjects who were assisting at the race but not running.

Interventions: Not applicable.

Main Outcome Measures: Overall pain and pain ratings on a pressure pain test before and after the event.

Results: Mean overall pain ± standard deviation on a 100-mm scale increased (P < .05) from 3 ± 6 mm before the run to 39 ± 28 mm after the run among the runners. The faster runners showed a mean reduction (P < .05) in pain ratings after the race of 15 ± 20 mm (on a 100-mm scale), whereas there was no change for the slower runners and controls. Findings were confirmed by model-based analysis.

Conclusions: The faster runners in a 100-mile (161-km) running race experience a modest temporary reduction in pressure pain perception that does not appear to be augmented by ongoing pain related to the exercise. The lack of a reduction in pain perception among the slower runners may be because an extreme bout of exercise of this nature can “exhaust” the systems responsible for exercise-induced analgesia in all but the most well-trained of runners, or that these systems were not activated because the slower runners were unable to maintain a high enough exercise intensity during the later stages of the race.

Key Words: Analgesia; Exercise; Pain; Pain threshold; Physical effort; Rehabilitation.

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

Previous research has shown that aerobic exercise can induce an acute analgesic effect. However, the duration and intensity of exercise required to elicit such an effect has not been fully clarified. Exercise-induced analgesia has been shown to occur from fairly high intensities for durations of as little as 8 to 10 minutes and up to 40 to 50 minutes, but longer bouts of exercise at low intensities have not been studied. Some work has suggested that intensities of over 70% of maximal aerobic capacity are required, and that pain thresholds increase with increasing intensities above this level. Our recent work showed that aerobic exercise of 30 minutes at 75% of maximal oxygen uptake (VO2max) resulted in a significant exercise analgesia effect, but 30 minutes of exercise at 50% of VO2max or 10 minutes at 75% of VO2max were inadequate to show the effect. This suggests that there may be thresholds for both intensity (>50%–70% of VO2max) and duration (>8–10 min) required for exercise-induced analgesia.

As far as we know, none of the studies examining postexercise analgesia have quantified the extent of naturally occurring pain as a result of the exercise. As such, no attention has been directed at how naturally occurring pain from the exercise might relate to the exercise-induced analgesic effect. An extreme bout of exercise, such as an ultramarathon competition, provides a venue to study pain perception in the presence of ongoing exercise-related pain. During and after such arduous exercise, some naturally occurring pain has been shown to be present. Whether the pain caused by this type of exercise would augment or reduce the usual exercise-induced analgesia that is observed after exercise is unknown.

The purpose of the present study was to determine whether pain perception is altered among competitors after completing a 100-mile (161-km) running race. Given the previous findings that exercise intensities of greater than 50% to 70% of VO2max were required to elicit exercise-induced analgesia and the recognition that this may be close to the intensity that is sustained for a 161-km run, we theorized that only the most trained runners, who were able to maintain a higher intensity throughout the race, would show exercise-induced analgesia. We also hypothesized that exercise-induced analgesia among the faster runners would be augmented by the presence of several hours of naturally occurring pain from the exercise.

METHODS

Participants

We performed the study at the Western States 100 Mile Endurance Run (June 25–26, 2005), a point-to-point trail run through the Sierra Nevada Mountains in Northern California. With over 5500 m of ascent and 6700 m of descent, this 161-km race is considered to be among the most arduous organized running events in the United States. The 2005 race was particularly challenging due to sections of snow cover on the trail for nearly the first 30 km. A 30-hour time limit is allowed for completion of the race.

Two groups of subjects participated in this study. Runners were recruited from the group of competitors in the race. Several weeks prior to the race, information about the study was sent to all entries. Some runners agreed to participate in the study in advance and others were recruited during race registration 1 to 2 days before the race. Control subjects were recruited among the crew members and race assistants who would be available 1 to 2 days before the race and again at the race finish, and had also indicated they would remain awake during the event.
Exclusion criteria for participation included the presence of any chronic pain condition, diseases affecting sensory nerves, pregnancy, the use of narcotic analgesics, and known cardiac, pulmonary, or metabolic disorders.

Procedures were approved by the institutional review board. All subjects were provided general information about the intent of the study in order to obtain informed consent, but the hypothesized results were not discussed in detail prior to completion of their participation in the testing.

Experimental Design
All subjects underwent 2 pain tests during the 2 days prior to the race. These 2 tests were separated by at least 30 minutes. The first test was considered to be for practice in order to allow the subjects to become familiar with the procedures. The second test was used in the data analysis.

After completing the race, the runners were directed to an area near the finish line where they underwent a third pain test. Control subjects who were assisting a runner in the study underwent the third pain test shortly before or after the runner finished, and those control subjects who were assisting with the race underwent the third test when they were available during the span of nearly 14 hours over which the runners were finishing.

Prior to the pain tests, we asked the subjects to list any pain medications taken in the previous 24 hours and to indicate their overall pain level by placing a mark on a 100-mm visual analog scale (VAS) anchored with “no pain” and “worst possible pain.”

Pressure Pain Testing
We performed the pain tests with a pressure pain stimulator similar to that previously used by others and the same device as in our previous work. The device uses a 6×0.25mm Lucite edge through which a constant force of 9.8N is applied against the dorsal surface of the middle phalanx of the nondominant index finger, halfway between the distal and proximal interphalangeal joints.

Subjects listened to a recorded message reviewing the procedures prior to each test. At 10-second intervals the audio recording instructed subjects to indicate the level of perceived pain by placing a mark on separate 100-mm VASs anchored with “no pain” on the left side and “worst possible pain” on the right side. Each pain stimulus lasted 2 minutes. During testing, the subjects were seated with arms supported on a table. Good reproducibility of these techniques with repeat testing at a 1-minute interval, as well as across days, has been previously shown.

Statistical Methodology
The main focus of the study was to examine for a change in pain perception as a result of running the race. This was a longitudinal study with 2 time points at which pain ratings were measured (before and after the race). Each pain test was also longitudinal in nature. However, the within-test pain ratings over time were not modeled, and were represented as a single summary (average) value. Each mark on the VAS was converted to a numeric value by measuring the distance in millimeters from the left side of the scale to the mark. Pain ratings were relatively stable during the last minute of the pain test, so the last 6 data points (ie, the final minute) were averaged to yield a single value for each trial.

Repeated before-after measurements induce correlation because they are shared by the same subject. This was taken into account using linear random-effects models for the analysis. The models allowed for fixed effects of explanatory variables as well as a random normally distributed subject-specific intercept term. Main effects as well as interactions were examined in search of the best model. The following variables were included in the model selection procedure: time, a binary variable modeling the effect of the race on within-subject outcome score; group, modeling the effect in runners compared with controls; a cluster of anthropomorphic and other subject characteristics including age (continuous), sex (binary), height (continuous), body mass (continuous), weekly running distance, training volume during the previous month (a measure of subject’s physical activity level), finish time (continuous and categorical dichotomized at the median) defined in runners and modeling the performance in the race, and delay (continuous) modeling the effect of delay in pain testing after the race. The search for the best model was not automatic. It involved testing key meaningful hypotheses based on a hierarchical family of models and the likelihood ratio test presented in table 1. A number of those hypotheses involved interaction effects targeting analgesic response operating differently in specific subgroups of subjects defined by the main variables mentioned above. These included interaction of group with time, which allowed us to study the effect of race separately in runners and controls, and time with finish time modeling the effect of race separately in slower and faster runners.

Observed versus predicted residual plots (not shown) were used to confirm model adequacy.

Data Analysis
There were 25 runners and 13 control subjects who agreed to participate in the study. Twenty-one of the runners completed the study, because 2 did not finish the race, one was in the medical tent for an extended period after completing the race, and one forgot to return for the pain test after the race. Eleven of the controls completed the study because 2 forgot to present for the final pain test. The time of day for the post-race pain test was distributed from approximately 5:00 to 11:00 AM for the controls, and approximately 12:30 to 11:30 AM for the runners. Two controls reported having obtained some sleep during the night.

We initially used descriptive statistics to summarize data by marginal means before and after the race, standard errors (SEs), and paired t tests targeting mean within subject change in pain test scores from before to after the race. Runners were split into 2 equally sized groups (slower, faster) by using a cutpoint at the median of finish time of 25.55 hours. A histogram of finish time is shown in figure 1.

The results shown in table 2 strongly suggest that the analgesic effect of the race predominantly manifests itself in the faster runners. The descriptive observation in table 2 was then confirmed in rigorous model-based analyses. A key sequence of models considered in the search and associated hypotheses is given in table 1. The search for the best model started with a model having 3 fixed effects modeled by categorical variables: time (post-race vs pre-race); group (runners vs controls), and time by group (interaction term modifying post- vs pre-race difference in runners as compared with the controls). The above variables directly pertain to the effect of exercise on analgesia, the main hypothesis of this study.

A round of forward selection attempted to include age (continuous), sex (categorical), height (continuous), body mass (continuous), weekly running distance (continuous), and finish time (continuous) in the initial time plus group plus time by...
EXERCISE ANALGESIA FROM AN ULTRAMARATHON, Hoffman

The number of participants in the study was 25.55 hours.

The median finish time of the subjects participating in the study was 1044 EXERCISE ANALGESIA FROM AN ULTRAMARATHON, Hoffman

The histogram of finish times for all competitors in the event and the subjects participating in the study. Median finish time of the participants in the study was 25.55 hours.

Table 1: Hypotheses Tested in Search for the Best Model

<table>
<thead>
<tr>
<th>Model Symbolic Representation</th>
<th>AIC Difference</th>
<th>Hypothesis (Term Removed)</th>
<th>LR</th>
<th>P*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group + time + group × time</td>
<td>1.55</td>
<td>No effect of race in controls (time, main effect) (group, main effect)</td>
<td>0.00</td>
<td>.965</td>
</tr>
<tr>
<td>Group + time × group</td>
<td>0.42</td>
<td>No difference between runners and controls pre-race (group, main effect)</td>
<td>1.58</td>
<td>.228</td>
</tr>
<tr>
<td>Time × group + age</td>
<td>1.49</td>
<td>No effect of age (age)</td>
<td>0.51</td>
<td>.474</td>
</tr>
<tr>
<td>Time × group + sex</td>
<td>1.35</td>
<td>No effect of sex (sex)</td>
<td>0.65</td>
<td>.419</td>
</tr>
<tr>
<td>Time × group + height</td>
<td>1.64</td>
<td>No effect of height (height)</td>
<td>0.36</td>
<td>.549</td>
</tr>
<tr>
<td>Time × group + mass</td>
<td>0.81</td>
<td>No effect of body mass (mass)</td>
<td>1.17</td>
<td>.276</td>
</tr>
<tr>
<td>Time × group + weekly running distance</td>
<td>1.02</td>
<td>No effect of exercise training volume before the race (weekly running distance)</td>
<td>0.98</td>
<td>.321</td>
</tr>
<tr>
<td>Time × group × finish time = intercept + I(time = post-race &amp; faster runners) + I(time = post-race &amp; slower runners)</td>
<td>2.00</td>
<td>No difference between slower runners and controls (pre-race &amp; faster runners)</td>
<td>&lt;0.00</td>
<td>.984</td>
</tr>
<tr>
<td>Time × group × finish time + delay = intercept + I(time = post-race &amp; faster runner) + delay</td>
<td>1.38</td>
<td>No effect of a delay in pain testing after the race (delay)</td>
<td>0.62</td>
<td>.431</td>
</tr>
</tbody>
</table>

NOTE. I(A)=1 if A is true, and I(A)=0 otherwise.
Abbreviations: AIC, Akaike information criterion; LR, likelihood ratio test statistics.

*P value for the hypothesis.

Having completed the above exploratory stage of best model selection, we refined the model to allow fine measurement of performance level by introducing finish time as a modifier for the effect of the race. Because continuous finish time showed no significance, we treated finish time as a categorical variable in the subsequent analysis. We were testing whether a subgroup of subjects with better performances showed a significant effect of the race. The target group was defined based on the finish time variable dichotomized using a cutpoint at the median. When analyzing a model containing the effect of time in runners (a group by time interaction), runners showed a mean ± SE within-subject reduction of pain rating of 6.6±4.0mm. This effect showed a trend toward significance at P equal to .109 and we followed with a targeted analysis of pain rating effects in runners.

Post hoc analysis using variable cutpoints revealed the lowest (unadjusted) P value of .008 at a cutpoint of 26 hours with the estimated effect ± SE of 13.8±4.9. Using this cutpoint, the mean finish times ± standard deviation (SD) were 22.8±2.1 and 28.5±1.4 hours for the faster and slower runners, respectively. Pain ratings among the faster runners were reduced from 52±20 to 37±18mm (fig 2).

The ongoing exercise-related pain showed no significant correlation with analgesia in the runners (P=.50) when the overall pain level was added to the model of race-induced changes in pain ratings.

The faster runners included 8 men and 4 women, and the slower runners included 8 men and 1 woman. Examination of some mutually related variables shed light on the difference in characteristics between the faster and slower runners (table 5). Faster runners tended to be younger (P=.053) and ran greater (P=.001) distances in training during the month before the race. At the same time, there were no differences in body mass index (BMI) (P=.080), pre-race pain ratings (52±20 vs 51±30, P=.92) (see fig 2), pre-race overall pain levels (1±3 vs 3±4, P=.30) (fig 3), post-race overall pain levels (39±28 vs 40±27, P=.94) (see fig 3), or time interval between finish-
ing the race and performance of the last pain test (35 ± 4.4 vs 17 ± 7 min, \( P = .25 \)). This latter observation was confirmed in model-based analysis including the variable delay (see table 1).

The control subjects included 5 men and 6 women. This group was of similar age and BMI to the runners, but was running less (\( P = .036 \) vs slower runners; \( P < .001 \) vs faster runners) than the runners (see table 3). Compared with the runners, the controls had no difference in pre-race pain ratings (42 ± 17, \( P = .25 \)) (see fig 2) and pre-race overall pain levels (5 ± 7, \( P = .16 \)) (see fig 3). Model-based analysis further showed no difference in pain ratings between runners and controls at baseline (main effect of group, \( P = .23 \)) (see table 1). Additionally, there was no change in pain rating (\( P = .91 \)) (see fig 2) or in overall pain levels (\( P = .30 \)) (see fig 3) among the controls between the pre-race and post-race evaluations. The correlation of pre-race and post-race pain ratings for the control group was highly significant (\( r^2 = .871, P < .001 \)).

The use of anti-inflammatory medications and acetaminophen was common among the runners. Of the 12 faster runners, one (8%) reported using anti-inflammatory medication within 6 hours of the pre-race test and 9 (75%) reported taking pain medication (anti-inflammatory medication and/or acetaminophen) that may have been within 6 hours of the post-race test. Among the 9 slower runners, 3 (33%) indicated they had taken anti-inflammatory medication within 6 hours of the pre-race test, and 8 (89%) indicated they had taken pain medication (anti-inflammatory medication and/or acetaminophen) that may have been within 6 hours of the post-race test. Timing of the pain medication prior to the post-race test was not always clear due to the difficulties some runners had in remembering such information after completion of the race. Control subjects reported no use of pain medications within 6 hours of the pain tests.

## RESULTS

Our model selection procedure described in the Data Analysis section converged at the final model that included separate effects of race (time) in slow and fast runners. The group of faster runners was defined prior to the analysis by placing a cutoff at 25.55 hours, the median of finish time. A significant effect of time in faster runners was identified (\( P = .017 \)) (table 4). No other effects were significant.

## DISCUSSION

The present study shows that faster runners perceived a painful stimulus to the finger to be less painful shortly after finishing a 161-km running race than before the race. The most interesting aspect of this study is that the exercise-induced analgesia effect was only present for the faster runners, and was not observed among the slower runners.

As might be expected, the faster runners tended to be younger and ran greater distances in training than the slower runners. Otherwise, the faster and slower groups of runners appeared similar in various characteristics including BMI, baseline pain ratings to the pain test, and use of pain medications during the race. Both groups also had similar increases in overall pain levels after the race. These mean overall pain levels of 39 to 40mm on a 100-mm VAS were comparable to the post-race "muscle soreness" values of about 4 to 5 on a 10-point Likert scale determined from another group of runners after the same event.9 Presumably, overall pain levels would have been even higher had they been measured during rather than after the race.

Activation of the endogenous opioid system is a mechanism that has been proposed to be involved in exercise-induced analgesia.1,2 Other mechanisms have also been suggested, such as activation of large afferent fibers incorporating the gate-control theory of pain, enhanced psychologic well-being, distraction, and activation of endogenous cannabinoids.7,17,18 Yet, even in the situation where there is ongoing pain, the evidence supports the involvement of the endogenous opioid system. Two animal studies have examined exercise-induced analgesia under such conditions. One study used an acute tissue injury model in mice19 and the other study used a chronic muscle pain model in rats.20 Both showed the presence of exercise-induced analgesia that was mediated through the endogenous opioid system.

It has been suggested that dysfunction of the endogenous opioid system may play a role in chronic pain. An impairment of the endogenous opioid antinociceptive system may contribute to the elevated pain sensitivity often observed among more disabled chronic pain patients.21 Yet, among less disabled patients, it has been suggested that chronic pain may serve as a primer producing up-regulated opioid antinociceptive responses to acute pain.22 The modest reduction in pain perception observed in the present study among the faster runners, which was comparable to what we have previously observed from moderate exercise,8,11 suggests that there was no augmentation of the opioid antinociceptive response from the ongoing pain associated with the run. Because the post-race overall pain levels were the same for the faster and slower runners, it would seem that the stimulus to the antinociceptive system due to naturally occurring pain from the exercise was similar between groups. Yet, we cannot be certain that the endogenous opioid antinociceptive system responded similarly in the 2 groups. Nonetheless, a more plausible explanation for the lack of evidence for exercise-induced analgesia in the slower runners might be that these runners were exercising at a lower intensity during the later stages of the race.

We have no data on the exercise intensity at which our subjects were working throughout the competition, but limited literature on this topic suggests that the percentage of \( V_{\text{O}_2}\text{max} \) that can be sustained for the duration of a 161-km ultramarathon is on the order of 50% of \( V_{\text{O}_2}\text{max} \).10,11 It would be anticipated that the slower runners tended to exercise at a lower percentage of \( V_{\text{O}_2}\text{max} \) compared with the faster runners due to greater limitations on their performance from general fatigue. This could be particularly likely during the later stages of the competition. Given that exercise intensities of over 50% to 70% of \( V_{\text{O}_2}\text{max} \) may be required to elicit a measurable

### Table 2: Descriptive Statistics for Pain Ratings Based on a Cutpoint of Median Finish Time for the Entire Group of Runners

<table>
<thead>
<tr>
<th>Group and Time</th>
<th>Marginal Mean ± SE</th>
<th>Pre vs Post</th>
<th>Mean ± SE of the Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-Race</td>
<td>Post-Race</td>
<td>t</td>
</tr>
<tr>
<td>Faster runners</td>
<td>53.51 ± 6.44</td>
<td>39.12 ± 5.72</td>
<td>14.37 ± 6.42</td>
</tr>
<tr>
<td>Slower runners</td>
<td>50.34 ± 9.42</td>
<td>49.80 ± 10.11</td>
<td>0.54 ± 7.38</td>
</tr>
<tr>
<td>Controls</td>
<td>41.97 ± 5.36</td>
<td>41.73 ± 5.99</td>
<td>0.24 ± 2.17</td>
</tr>
</tbody>
</table>

Arch Phys Med Rehabil Vol 88, August 2007
exercise-induced analgesic effect, it is conceivable that the exercise intensity was inadequate for the slower runners to induce an exercise-analgesic effect.

The present study was not directed at clarifying the underlying mechanism of exercise-induced analgesia. Still, it is interesting to examine the present results in view of some previous work done at the Western States 100 Mile Endurance Run. Bortz et al reported the findings of elevations in β-endorphin concentrations in the blood at 97km and at the finish among competitors in the 1980 race. Although the subjects who were tested at the 2 points on the course were not the same and the authors did not comment on the relative performance level of the subjects who participated in their study, the β-endorphin concentrations appeared to be lower at the finish than at 97km. This could represent a reduction in β-endorphin production during the later stages of the race simply from “exhaustion” of this physiologic function, or due to a reduction in exercise intensity to the point that the endogenous opioid system was not fully activated in some runners.

Study Limitations
There is always a concern about reproducibility of a measurement that involves some subjectivity as is the case with the pressure pain test that was used in this study. Our previous work has shown acceptable reproducibility for repeat pain tests on a given day after a single practice test on a previous day, as well as for repeat pain tests across days. As such, we believe that our design that involved 2 pain tests before the race, with the first serving as a practice, was adequate to assure acceptable reproducibility of the measurement.

Table 3: Selected Characteristics of the Subjects

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Faster Runners</th>
<th>Slower Runners</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n=12; 8 men, 4 women)</td>
<td>(n=9; 8 men, 1 woman)</td>
<td>(n=9; 8 men, 1 woman)</td>
</tr>
<tr>
<td>Age (y)</td>
<td>42±6</td>
<td>49±9</td>
<td>44±11</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>22.3±2.0</td>
<td>24.3±2.6</td>
<td>22.5±1.7</td>
</tr>
<tr>
<td>Weekly running distance during previous month (km)</td>
<td>120±31*</td>
<td>65±33*</td>
<td>36±21*</td>
</tr>
<tr>
<td>Finish time (h)</td>
<td>22.8±2.1*</td>
<td>28.5±1.4*</td>
<td>NA</td>
</tr>
</tbody>
</table>

NOTE. Values are mean ± SD. Abbreviation: NA, not applicable. *Statistical differences between groups were identified with unpaired t tests.

**Fig 2.** Mean pain ratings for (A) faster runners, (B) slower runners, and (C) control subjects before (○) and after (●) the competition. *Pain ratings averaged over the last minute of the test were significantly different (P<.05) than the pre-race test. Error bars represent 1 SD, and for clarity are displayed in only 1 direction for each test.

**Fig 3.** Overall pain levels for the 3 groups of subjects. *Post-race pain ratings were significantly different (P<.05) than pre-race ratings for the 2 runner groups, and post-race pain ratings differed significantly between the runners and controls.
reproducibility. Indeed, the lack of change in pain ratings among the control subjects, in which no alteration would be expected, provides further support for the reproducibility of the pain test after a single practice test.

It should be noted that we were vigilant in recruiting a control group that included subjects of similar characteristics to the runners. Although it turned out that the control group included a higher percentage of female subjects and was not training at the same level as the runners, there is no reason to suspect that there would have been a variation in pain ratings across tests had there been more men or more active subjects in the control group. Most importantly, the majority of the control subjects were busy and awake throughout the duration of the race. This shows that the reduction in pain ratings among the faster runners was not simply a matter of the stress from having been awake for nearly 24 hours or longer.

Field studies are often fraught with variables that are difficult to control. Like other field studies, this one has some potentially confounding issues that warrant comment. Given that this study was performed around a challenging competition for which these athletes had invested considerable effort in order to participate, it was necessary for the study to avoid significant intrusions. As such, we were unable to control the use of pain medications before and during the race. In fact, 4 of the runners reported using anti-inflammatory medications within 6 hours before undergoing the pre-race pain test, and all but 3 of the runners reported taking a pain medication (acetaminophen and/or anti-inflammatory drug) sometime during the race. In contrast, none of the control subjects reported having used a pain medication within 6 hours of any of the pain tests. Because the use of pain medication before (1/12 in the faster group, 3/9 in the slower group) and after (9/12 in the faster group, 8/9 in the slower group) the race appeared comparable between the faster and slower runners, we do not think that this factor can account for the finding of a different effect for the 2 groups of runners.

It was also important that none of the subjects exercised shortly before the pre-race pain tests, and that the control subjects did not exercise shortly before the post-race pain tests. Each subject was queried before each pain test, and we found that there was no significant exercise shortly prior to pain testing that should have confounded the results.

One variable that we were unable to control to our desired level was the time interval from finishing the race to the post-race pain test. Three of the faster runners were not willing to undergo the post-race pain test until 0.85 to 2.78 hours after finishing; whereas all of the slower runners were tested within the desired interval. Nonetheless, given that the exercise-induced analgesic effect is likely to be less evident the longer it has been after the completion of exercise, this issue is not thought to account for the difference between the faster and slower runners. When excluding the 3 runners showing more than a 30-minute interval from finishing to post-race pain testing, the effect of the race was no longer significant (P=.125). However, the estimate of the effect in faster runners (−11.51mm) still points at a strong trend toward increased analgesia. Therefore, we attribute the loss of significance to a reduction in sample size. Furthermore, in the model-based analysis the delay variable showed no significance (P=.51).

Another potentially concerning issue relates to differences in air temperature at the time of the post-race testing. Because the faster runners finished between 12:00 PM and 7:00 AM, the air temperatures were lower (range, 12°–16°C) than when the slower runners finished (range, 16°–21°C). It is recognized that cooling slows sensory nerve conduction velocities, although it is not known if this amount of variation in air temperature would affect pain ratings in the type of test we performed. Nonetheless, because nearly half of the controls performed the last pain test within the temperature range of the faster runners and the other half performed the test within the temperature range of the slower runners, and there was no apparent relationship of pain ratings with air temperature among the controls, we have discarded this issue as having affected the results.

We were concerned that some of the runners might have swelling in the fingers that could alter the pain ratings during the post-race pain test. Some finger swelling is not unusual among competitors in endurance events due to various factors including the development of low serum sodium levels. It is recognized that cooling slows sensory nerve conduction velocities, although it is not known if this amount of variation in air temperature would affect pain ratings in the type of test we performed. In fact, 4 of the runners reported taking a pain medication (acetaminophen and/or anti-inflammatory drug) sometime during the race. In contrast, none of the control subjects reported having used a pain medication within 6 hours of any of the pain tests. Because the use of pain medication before (1/12 in the faster group, 3/9 in the slower group) and after (9/12 in the faster group, 8/9 in the slower group) the race appeared comparable between the faster and slower runners, we do not think that this factor can account for the finding of a different effect for the 2 groups of runners.

It was also important that none of the subjects exercised shortly before the pre-race pain tests, and that the control subjects did not exercise shortly before the post-race pain tests. Each subject was queried before each pain test, and we found that there was no significant exercise shortly prior to pain testing that should have confounded the results.

One variable that we were unable to control to our desired level was the time interval from finishing the race to the post-race pain test. Three of the faster runners were not willing to undergo the post-race pain test until 0.85 to 2.78 hours after finishing; whereas all of the slower runners were tested within the desired interval. Nonetheless, given that the exercise-induced analgesic effect is likely to be less evident the longer it has been after the completion of exercise, this issue is not thought to account for the difference between the faster and slower runners. When excluding the 3 runners showing more than a 30-minute interval from finishing to post-race pain testing, the effect of the race was no longer significant (P=.125). However, the estimate of the effect in faster runners (−11.51mm) still points at a strong trend toward increased analgesia. Therefore, we attribute the loss of significance to a reduction in sample size. Furthermore, in the model-based analysis the delay variable showed no significance (P=.51).

Another potentially concerning issue relates to differences in air temperature at the time of the post-race testing. Because the faster runners finished between 12:00 PM and 7:00 AM, the air temperatures were lower (range, 12°–16°C) than when the slower runners finished (range, 16°–21°C). It is recognized that cooling slows sensory nerve conduction velocities, although it is not known if this amount of variation in air temperature would affect pain ratings in the type of test we performed. Nonetheless, because nearly half of the controls performed the last pain test within the temperature range of the faster runners and the other half performed the test within the temperature range of the slower runners, and there was no apparent relationship of pain ratings with air temperature among the controls, we have discarded this issue as having affected the results.

We were concerned that some of the runners might have swelling in the fingers that could alter the pain ratings during the post-race pain test. Some finger swelling is not unusual among competitors in endurance events due to various factors including the development of low serum sodium levels. If finger swelling was present, it could, in effect, cause some “cushioning” between the edge of the pressure pain stimulator and the finger. This could either lower the pain ratings or cause a shifting of the pain rating–time curve to the right. However, we are unable to formulate a plausible explanation for why this effect would be greater among the faster runners than the slower runners. Thus, we do not believe that this factor accounts for the reduced pain ratings that were observed among the faster runners.

Small sample size is recognized as an important limitation in this study. Important effects may have been masked because the study was only powered to detect large differences. It is quite likely that we would have discovered a continuous race effect modification by finishing time in a larger study, as well as other predictors of the association between exercise intensity and analgesia.

CONCLUSIONS

Fastest runners in a 161-km running race experience a modest temporary reduction in pressure pain perception. The magnitude of this analgesic effect was comparable to that previously observed after modest exercise, without evidence for an augmented response due to ongoing pain from the exercise. It is unknown whether the lack of a reduction in pain perception among the slower runners was because an extreme bout of exercise. It is suspected that some of the runners might have swelling in the fingers that could alter the pain ratings during the post-race pain test. Some finger swelling is not unusual among competitors in endurance events due to various factors including the development of low serum sodium levels. If finger swelling was present, it could, in effect, cause some “cushioning” between the edge of the pressure pain stimulator and the finger. This could either lower the pain ratings or cause a shifting of the pain rating–time curve to the right. However, we are unable to formulate a plausible explanation for why this effect would be greater among the faster runners than the slower runners. Thus, we do not believe that this factor accounts for the reduced pain ratings that were observed among the faster runners.

Small sample size is recognized as an important limitation in this study. Important effects may have been masked because the study was only powered to detect large differences. It is quite likely that we would have discovered a continuous race effect modification by finishing time in a larger study, as well as other predictors of the association between exercise intensity and analgesia.

Acknowledgment: We thank the leadership of the Western States Endurance Run for support of the research.

References

Learning Effects of Repetitive Administrations of the Sensory Organization Test in Healthy Young Adults

Diane M. Wrisley, PhD, PT, NCS, Marilee J. Stephens, PhD, Shaun Mosley, PT, DPT, Adam Wojnowski, AuD, Jordan Duffy, PT, DPT, Robert Burkard, PhD


Objectives: To evaluate the learning effect of multiple administrations of the Sensory Organization Test (SOT) on performance and to begin to establish clinical meaningful change scores for the SOT.

Design: Descriptive case series.

Setting: University-affiliated clinic.

Participants: Healthy young adults (6 men, 7 women; mean age 24±2.5y).

Intervention: All subjects performed the standardized SOT using the SMART EquiTest 5 times over a 2-week period, and 1 month later.

Main Outcome Measure: Composite and individual SOT test condition standardized equilibrium scores.

Results: Test-retest reliability (intraclass correlation coefficient model 2.3) of the composite (.67) and equilibrium score (range, .35-.79) were fair to good. Repeated-measures analysis of variance revealed a significant (P<.05) increase in the composite and equilibrium scores for conditions 4, 5, and 6 over the 5 sessions that plateaued after the third session, and were retained at 1 month. The 95% confidence interval for the composite score change from session 1 to session 4, the plateau of the learning effect, was 3.9 to 8.1.

Conclusions: Although the findings of this study would indicate that multiple baseline measures are desirable for the more challenging conditions, a composite change of greater than 8 points would indicate change due to rehabilitation.

Key Words: Balance; Learning; Posture; Rehabilitation; Reproducibility of results.

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

Human postural control has been studied in both healthy subjects and patients with a variety of pathologic conditions to understand the neural mechanisms (sensory, motor) involved in a relatively automatic, yet complex, task, and to develop training and rehabilitation tools to avoid or correct deficits in the system. One means of taxing the postural control system is through changing the characteristics of the support surface. Platform perturbations allow for the investigation of the support surface. Platform perturbations allow for the investigation of both the sensory influences and interactions on balance control and the motor outcomes. Deviations in performance seen in those with pathologic conditions has allowed the development of rehabilitation techniques in order to increase stability in these populations and therefore decrease the likelihood of falls.

The ability to maintain an upright posture within a given sensory environment (balance), is dependent on sensory information to detect the position and movement of the body so that appropriate movement responses can be generated. The Sensory Organization Test (SOT) of Computerized Dynamic Posturography was developed in order to identify the relative contribution of the 3 main sensory systems involved in balance (vision, vestibular, somatosensory). The test attempts to isolate the various sensory contributions by either removing or distorting (via sway-referencing the visual surround or the surface platform) the visual and/or somatosensory inputs to the postural-control system. The resultant 6 conditions progress from the most stable (eyes open, solid support surface) to the least stable (sway-referenced vision and surface). The 6 conditions of the SOT are described in table 1. Moderate to good test-retest reliability for the SOT over 2 sessions 1 week apart has been shown.

Scores on the SOT allow the differentiation between subjects with normal and abnormal vestibular function, as well as identifying older adults at risk for falls.

The SOT paradigm has been used to show the altered use of sensory input in multiple populations, including older adults, young children, and people with vestibular dysfunction, peripheral neuropathy, and Parkinson’s disease. Clinically, several studies have used this protocol pre- and postintervention during different treatment scenarios. Improved SOT scores have been shown in patients with Parkinson’s disease and osteoporosis after balance training, and in patients with central and peripheral vestibular dysfunction after vestibular rehabilitation.

However, it has been reported that repeated exposure to a given perturbation of the postural control system allows for learning of more efficient postural strategies to maintain balance both within a session and over time. Even in simple quiet stance with eyes closed, a progressive reduction in sway area and sway path was seen with repetition, with the authors suggesting that the body learns to move to a “safer” position that results in less energy expenditure by reducing the amount of sway. This has also been shown in more complex balance tasks, such as standing with a narrow base of support, or with vibration to the ankle tendons. If repeated exposure to given perturbations leads to an improvement in postural stability, this may serve as a confound to the use of the SOT to measure improvement in balance in intervention studies.
Table 1: The 6 Conditions of the SOT

<table>
<thead>
<tr>
<th>Condition</th>
<th>Vision</th>
<th>Surface</th>
<th>Visual Surround</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Eyes open</td>
<td>Stable</td>
<td>Stable</td>
</tr>
<tr>
<td>2</td>
<td>Eyes closed</td>
<td>Stable</td>
<td>Stable</td>
</tr>
<tr>
<td>3</td>
<td>Eyes open</td>
<td>Stable</td>
<td>Sway-referenced*</td>
</tr>
<tr>
<td>4</td>
<td>Eyes open</td>
<td>Sway-referenced*</td>
<td>Stable</td>
</tr>
<tr>
<td>5</td>
<td>Eyes closed</td>
<td>Sway-referenced*</td>
<td>Stable</td>
</tr>
<tr>
<td>6</td>
<td>Eyes open</td>
<td>Sway-referenced*</td>
<td>Sway-referenced*</td>
</tr>
</tbody>
</table>

*The surface or the surround moves at a gain of 1 with center of pressure movements so that the visual field or the ankle angles stay relatively constant.

Ford-Smith et al evaluated the test-retest reliability of the SOT in noninstitutionalized older adults. They found moderate test-retest reliability (intraclass correlation coefficient [ICC], .66 for composite score) when 2 sessions of the test were administered 1 week apart. The ICC for the individual conditions ranged from .26 to .68 with improvements from the first to second session of up to 18%. Tsang et al found moderate to good reliability of the individual SOT conditions with ICC scores ranging from .72 to .93. No reliability statistics were presented for the composite score. This moderate reliability of individual conditions and the composite score illustrates that additional study is needed to identify the true test-retest reliability of the SOT, and to differentiate the normal learning of the task over time from improvements in balance due to interventions. Therefore, the purpose of this study was to see what learning effects are seen in SOT scores over multiple test sessions in healthy young adults and to determine if this improvement is retained over time. An additional purpose was to begin to establish a framework on which the clinically significant change in SOT scores could be determined.

**METHODS**

**Participants**

We recruited 13 young healthy subjects (6 men, 7 women; mean age, 24±4y) for this study. Young adults were included in this study because we anticipated that they would show the smallest learning effect and would allow us to estimate the minimum change necessary for determining change due to rehabilitation. The sample size of 13 participants provide a power of .95 for detecting a large effect size (F=40) using repeated-measures analysis of variance (ANOVA) with a correlation of 0.5 between sessions. All participants were university students ranging in age from 21 to 36 years old. All subjects met the following inclusion criteria: (1) no reported history of neurologic, orthopedic, or muscular injury, (2) normal bilateral lower-extremity strength assessed via a manual muscle test, (3) able to stand independently for 20 minutes, (4) able to stand on toes and heels, and have normal functional range of motion in ankles, knees, and hips, (5) have normal somatosensory function in the feet as measured by the Weinstein Enhanced Sensory Test and vibration sensation at 128Hz, (6) no history of neck injury, whiplash, or current complaints of neck pain, and (7) not currently on any medical treatment that might affect balance. Subjects were asked to refrain from alcohol consumption for 48 hours before any test session and alcohol consumption was monitored through subject self-report. The study was approved by the Health Sciences Institutional Review Board and informed consent was obtained from each person. All subjects completed a general health questionnaire, the Dizziness Handicap Inventory (DHI) and the Activities-specific Balance Confidence (ABC) Scale to ensure that they met the inclusion criteria. The DHI is a 25-item questionnaire that quantifies a person’s perceived handicap due to dizziness. It is scored on a 0 to 100 scale with higher scores indicating greater perceived handicap. The DHI has high reliability and validity. All subjects included in this study scored a 0 on the DHI. The ABC is a 16-item questionnaire that requires the subject to rate how confident they are that they will not lose their balance on 16 tasks ranging from walking around the house to walking on icy sidewalks. The ABC is scored on a percentage scale, with 100% indicating that subjects are completely confident that they will not lose their balance and 0% indicating that they feel they will lose their balance during that task. ABC scores for the subjects participating in this study were all greater than 95%, with a mean of 98%±2%.

**Posturography Test Procedure**

All subjects completed the standardized SOT on the NeuroCom Smart EquiTest. The SOT consists of 6 sensory conditions (table 1). Subjects completed 3 trials for each of the 6 sensory conditions during each session. Subjects stood on the platform in bare feet with the feet placed 5.7cm apart and the medial malleus aligned with the axis of platform rotation. Foot position was marked on the platform to allow consistency between trials and sessions. Subjects wore a harness that attached overhead and prevented falls but did not limit sway. Subjects were asked to stand quietly with their arms across their chest and their eyes open or closed (depending on the condition). An examiner remained stationed behind each subject for safety throughout the test.

**Testing Procedure**

Each subject completed 5 testing sessions over a 2-week period with a retention test 1 month later (1 subject completed only the initial 5 testing sessions and did not return for the 1-mo follow-up session). Subjects had a mean of 1.9±1.1 days of rest between the 5 test sessions and a mean of 29.2±4.4 days between the test sessions and the retention test. Subjects completed the SOT according to the manufacturer’s instructions, with 3 repetitions of each condition in each session. Each trial lasted 20 seconds. The conditions were performed in subsequent order 1 through 6 for each of the 6 sessions. No subjects fell during the performance of any condition. A computer-generated equilibrium score was calculated for each trial using the formula provided by NeuroCom. The equilibrium score is a percentage that compares the subject’s anteroposterior center of pressure sway (in degrees) with the theoretical limits of stability, that is, maximum peak to peak sway of 12.5° (8° forward, 4.5° backward). The equilibrium scores range from 0% to 100%, with 100% indicating perfect stability and 0% indicating a fall. The 3 trials for each condition were averaged for an equilibrium score for each condition. A composite score, the mean of the average equilibrium score for all trials of conditions 1 and 2 and the 3 trials of conditions 3 through 6, was also computer generated based on a formula from NeuroCom for each session and was used for data analysis. Therefore, the composite score is a weighted average that emphasizes the more difficult balance conditions.

**Data Analysis**

We analyzed differences in equilibrium and composite scores across the 6 sessions using repeated-measures ANOVA. In order to perform post hoc testing and to determine the effect of trial, a multivariate analysis with trial and session as factors,
and subject and days of rest as co-factors was also performed. Post hoc tests included the Bonferroni test to determine differences between sessions and the Dunnett $t$ to determine which sessions were significantly greater than the first session. Significance level was set at $P < 0.05$. Test-retest reliability was quantified using intraclass correlation coefficient model 2,3 ($ICC_{2,3}$). ICC$_{2,3}$ was used as 3 trials of each condition were averaged for the individual condition and the composite scores. Using the Fleiss criteria, we define ICC values of less than 0.4 having poor reliability; 0.4 to 0.75 fair to good reliability, and scores above 0.75 as having excellent reliability. Paired $t$ tests were performed between scores for session 5 and 6 to determine if the learning effect was retained at 1 month. All analysis was completed using SPSS.

**RESULTS**

**Learning Effect**

A significant difference was found in equilibrium scores between all 3 trials in all sessions for conditions 4, 5, and 6 but not for conditions 1 through 3. All subjects achieved normal scores on the SOT and all subjects had an increase in their score, but not for conditions 1 through 3. Mean and standard deviations (SDs) for the equilibrium scores of individual conditions and the composite scores are listed in Table 2. On post hoc testing for condition 4 and 5, significant differences were found between sessions 1 and sessions 3 through 6, and sessions 2 and sessions 3 through 6; for condition 6, significant differences were found between session 1 and sessions 2 through 6, and for the composite score, significant differences were found between session 1 and sessions 2 through 6, and session 2 and all other sessions. Equilibrium scores were significantly greater in sessions 3 through 6 than in session 1 for conditions 4 and 5 and significantly greater in session 2 through 6 than in session 1 for condition 6 (Table 2). The composite score was significantly greater in sessions 2 through 6 than in session 1. To determine the amount of change necessary to indicate improvement beyond learning of the task, 95% CIs of the change from session 1 were calculated for each condition and the composite score (see Table 3) for each session. The upper level of the CI for the composite score change for session 4, the plateau of the learning effect, was 8.1, so an improvement of greater than 8 points on the composite score would be considered a change greater than the learning of the task.

**Test-Retest Reliability**

Fair to good test-retest reliability was found for the SOT composite score from session 1 to session 2 with an average of $0.47$±0.9 days in between testing was $0.67$ (ICC$_{2,3}$). Individual equilibrium scores for all conditions except condition 3 (stable surface, sway referenced vision) were also fair to good, with scores ranging from 0.43 to 0.79 from session 1 to session 2. Condition 3 showed poor test-retest reliability with an ICC$_{2,3}$ of 0.35.

**Retention**

The majority (10/12 [83%]) of subjects either retained or improved their performance on the SOT composite score when tested 1 month later (session 6). Five subjects retained their score, 5 subjects improved their score (average, 1.6 points) and 2 subjects had a decrease in score (average, 2 points). All composite SOT scores were greater in sessions 5 and 6 than in session 1. The subject and number of days rest between sessions as covariate yielded no statistical difference. No statistical differences between session 5 and 6 in individual conditions or composite score were found using paired $t$ tests.

**DISCUSSION**

This study showed, for the first time, the learning effect seen with repetitive SOT administrations. To account for this learning effect, it could be recommended that clinicians administer multiple baseline sessions to establish a steady performance

---

Table 2: Mean Equilibrium Scores* on Individual SOT Conditions and Composite Scores† for Each Session

<table>
<thead>
<tr>
<th>Condition</th>
<th>Session 1</th>
<th>Session 2</th>
<th>Session 3</th>
<th>Session 4</th>
<th>Session 5</th>
<th>Session 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Eyes open; firm surface (%)</td>
<td>95.3±1.6</td>
<td>95.4±1.3</td>
<td>95.9±1.2</td>
<td>95.8±0.8</td>
<td>95.8±1.1</td>
<td>95.6±1.0</td>
</tr>
<tr>
<td>2: Eyes closed; firm surface (%)</td>
<td>93.6±2.2</td>
<td>93.7±1.6</td>
<td>93.8±1.6</td>
<td>93.0±1.9</td>
<td>93.4±1.8</td>
<td>93.4±1.4</td>
</tr>
<tr>
<td>3: Sway reference vision; firm surface (%)</td>
<td>91.6±4.0</td>
<td>93.1±2.4</td>
<td>92.6±2.1</td>
<td>93.6±2.0</td>
<td>93.5±2.2</td>
<td>94.1±1.8</td>
</tr>
<tr>
<td>4: Eyes open; sway reference surface (%)</td>
<td>87.3±6.0</td>
<td>90.3±5.8</td>
<td>91.9±1.6</td>
<td>92.8±1.9</td>
<td>94.0±1.3</td>
<td>93.2±2.0</td>
</tr>
<tr>
<td>5: Eyes closed; sway reference surface (%)</td>
<td>74.6±3.6</td>
<td>77.5±4.5</td>
<td>81.0±3.4</td>
<td>83.6±3.1</td>
<td>83.3±2.7</td>
<td>83.3±3.9</td>
</tr>
<tr>
<td>6: Sway reference vision and surface (%)</td>
<td>72.9±7.1</td>
<td>79.6±6.1</td>
<td>82.1±4.7</td>
<td>82.8±5.6</td>
<td>84.7±4.1</td>
<td>85.1±3.4</td>
</tr>
<tr>
<td>Composite score</td>
<td>83.4±3.0</td>
<td>86.3±2.3</td>
<td>88.0±1.5</td>
<td>89.2±2.1</td>
<td>89.5±1.5</td>
<td>89.9±2.0</td>
</tr>
</tbody>
</table>

**NOTE.** Values are mean ± SD.
*Percentage of the subject’s anteroposterior center of pressure sway to the theoretical limits of stability.
†The mean of the average equilibrium score for all trials of condition 1 through 6.
prior to administering an intervention in order to document changes due to rehabilitation. However, the reality of clinical practice may not support the use of multiple baseline administrations and the use of SOT testing may act as training for patients with balance dysfunction. The administration of multiple baseline trials may also promote the learning effects shown in this study. Therefore, we have attempted to establish change values that would represent the amount attributable to learning. This is important, because it provides guidelines for interpreting the improvement in SOT scores after an intervention. Based on the results of this study we recommend using the criterion of 8 point improvement in the composite score for indicating improvement beyond the learning of the task. If previous outcome studies are reexamined using this criterion, we find greater support that balance function improved after intervention. Badke et al\textsuperscript{19} show a mean change of 8.5 points (from 42.6 to 51.1) in the composite score after vestibular rehabilitation in people with peripheral and central vestibular dysfunction. This change was not found to be statistically significant (\(P=.08\)), but using the criteria presented here it would show minimal change slightly beyond that expected for learning. Sinaki and Lynn\textsuperscript{18} found changes of 13 to 37 points in 3 women with osteoporosis after an exercise and posture training support intervention whereas the 2 women who received only the exercise intervention showed a change of \(-2\) and 3 points in their SOT scores.

The clinical relevance of this change score needs to be established. Further research is needed to determine the magnitude of change that would correlate with improvements in functional status. The magnitude of the learning effect may also be different for various populations. We expect that older adults will show a greater learning effect when performing the SOT and that they may require additional sessions for the learning effect to plateau.

Significant learning effects occurred for SOT conditions 4, 5, and 6 over the 5 sessions, whereas conditions 1 through 3 did not. Conditions 4 through 6 are more complex postural tasks and are thought to require vestibular information to maintain upright. Other studies have shown that with repeated exposure to a given perturbation or balance activity, performance on that task improves; this improvement is greatest in more complex tasks such as standing with narrow base of support and with sensory conflict such as vibration.\textsuperscript{20,22,24} Greater learning may have been seen in conditions 4 through 6 because of the complexity of the task, the novelty of the task, or the dependence of those tasks on vestibular information. It would be interesting to determine if the same level of learning is seen with people with visual or proprioceptive deficits or in people with highly trained vestibular systems (ie, ice skaters or dancers).

It has been suggested that at least 2 processes of adaptation occur with learning a novel postural task.\textsuperscript{20,25} One is short-term adaptation that reduces the postural sway by either changing the postural strategies (ie, increasing the stiffness in the ankles) or through reweighting of sensory information. Short-term adaptation was observed in this study with the equilibrium scores of individual trials improving within each session. Another form of adaptation is the long-term process; the profound effect previous experience of the postural task has on the development of a strategy for maintaining postural control.\textsuperscript{25,33,34} The majority of the subjects in this study retained or improved their composite scores on the 1 month follow-up session. The fact that 42\% (5/12) of the subjects improved their performance after a 1-month rest indicates further adaptation.

The SOT composite and equilibrium scores showed fair to good test-retest reliability according to the criteria by Fleiss.\textsuperscript{32}
when tested in healthy young adults with 1 to 3 days between tests. This is in agreement with the Ford-Smith et al. results from repeated administrations of the test 1 week apart in noninstitutionalized older adults, but is lower than the test-retest reliability found by Tsang et al. when testing 12 older adults 1 week apart. Surprisingly, the 2 conditions with the highest test-retest reliability are the hardest conditions for subjects to perform and have the greatest variability. The ICCs for the other conditions may have been lower due to a lack of variability. Although not ideal, this level of test-retest reliability has been considered acceptable for clinical tools. When used in combination with the improvement criteria discussed above, this reliability should allow for accurate interpretation of the SOT scores.

The test-retest reliability of the SOT composite score between sessions 4 and 5 and sessions 5 and 6 were excellent, further substantiating that the plateau in performance of the SOT occurs after the fourth administration of the test and that the improvements are retained even after a month. Variability was seen across trials and testing sessions, most likely due to learning. This learning effect manifested itself as lower standard deviations for conditions 4 through 6 across sessions and in lower ICC values from session 1 to 2 compared with the ICCs from session 4 and 5 or 5 and 6.

Study Limitations

One limitation of this study is that it was performed with healthy young adults. It is interesting that even active young adults can improve their balance function on this clinical assessment tool. However, the study findings need to be expanded to other groups. Older adults and people with sensory and motor pathologies may show a different learning effect. Additional studies are also needed to determine the sensitivity and specificity of the criteria introduced.

CONCLUSIONS

Healthy young adults show a learning effect when provided with 5 repetitions of the SOT over a 2-week period. The learning effect appears to plateau around session 3 and 4, and occurs primarily in the composite score and conditions 4 through 6. Multiple baseline measures of the SOT should be administered in order to document change due to rehabilitation. As an alternative, improvements of more than 8 points in the composite score indicate recovery beyond the effect of adaptation to the SOT itself. This improvement criterion provides an additional method for assessing outcomes in intervention studies.

References


Suppliers
b. Version 13; SPSS Inc, 233 S Wacker Dr, 11th FL, Chicago, IL 60606.
Strength Training Improves the Tri-Digit Finger-Pinch Force Control of Older Adults

Justin W. Keogh, PhD, Steve Morrison, PhD, Rod Barrett, PhD


Objective: To investigate the effect of unilateral upper-limb strength training on the finger-pinch force control of older men.

Design: Pretest and post-test 6-week intervention study.

Setting: Exercise science research laboratory.

Participants: Eleven neurologically fit older men (age range, 70–80y).

Intervention: The strength training group (n=7) trained twice a week for 6 weeks, performing dumbbell bicep curls, wrist flexions, and wrist extensions, while the control group subjects (n=4) maintained their normal activities.

Main Outcome Measures: Changes in force variability, targeting error, peak power frequency, proportional power, sample entropy, digit force sharing, and coupling relations were assessed during a series of finger-pinch tasks. These tasks involved maintaining a constant or sinusoidal force output at 20% and 40% of each subject’s maximum voluntary contraction. All participants performed the finger-pinch tasks with both the preferred and nonpreferred limbs.

Results: Analysis of covariance for between-group change scores indicated that the strength training group (trained limb) experienced significantly greater reductions in finger-pinch force variability and targeting error, as well as significantly greater increases in finger-pinch force, sample entropy, bicep curl, and wrist flexion strength than did the control group.

Conclusions: A nonspecific upper-limb strength-training program may improve the finger-pinch force control of older men.

Key Words: Aging; Exercise therapy; Grasp; Hand strength; Rehabilitation.

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

ACTIVITIES OF DAILY LIVING (ADLs) that require the dexterous manipulation of hand-held objects, such as drinking, eating, and writing, typically deteriorate as a consequence of aging. Although hand function is multifactorial in nature, an optimal level of isometric force control appears to be particularly important. The reduced hand function of older adults may therefore reflect, at least in part, the age-related loss of isometric finger-pinch force control.

The age-related reduction in muscular strength and neuromuscular function, which includes changes at the level of the motor unit, is an obvious contributing factor to the decline in force control in older people. Specific changes in motor unit properties associated with the aging process include altered motor unit size and firing rate, as well as altered relations between motor unit recruitment and motor unit firing rate. Age-related changes in the power spectral, regularity, sharing, and coupling of the digit forces have also been observed and may also predispose older adults to a loss of digit force control. The digit force outputs of older adults tend to have significantly more power at low frequencies and/or are more regular than those of young adults. Further, older adults have a greater percentage contribution (force share) of the lateral digits and a reduced medial digit(s) force share, as well as lower levels of digit force coupling than do young adults during 5- and tri-digit finger-pinch tasks, respectively.

Resistance training can produce substantial neural and coordinative adaptations in both young and older adults. Of particular relevance to digit force control, resistance training can significantly improve strength, intra- and intermuscular coordination, and can reduce motor unit firing rate variability. Consequently, it is not surprising that resistance training can produce significant improvements in the isometric digit force control of older adults. In many of the cited studies, the improvements in force control were observed during tasks that were similar to the exercises performed in training (eg, index finger abduction). It has yet to be determined whether there would be changes in force control in more general movement tasks that were not practiced during training.

The generality of response is further clouded by the fact that the greatest improvements in older adults’ digit force control after training have occurred in the trained limb when producing constant forces no greater than 20% of maximum voluntary contraction (MVC). Consequently, it remains unclear whether there are the same improvements in digit force control after the performance of tasks with greater MVC requirements, or with time varying target forces. We can gain a greater insight into the generality of the training response by assessing the change in force control in both the trained and untrained limbs across a range of task constraints, that is, at a variety of force levels and with various target shapes (ie, sinusoidal as well as constant forces) and with tasks involving multiple fingers.

We conducted this study to determine whether a general (nonspecific) unilateral upper-limb strength-training program could significantly improve finger-pinch force control in older adults. We assessed changes in force control in both trained and untrained limbs with 4 tri-digit finger-pinch conditions that reflected all combinations of mean force level (20% MVC and 40% MVC) and target shape (constant and sinusoidal).
hypothesized that strength training would significantly reduce finger-pinch force variability and targeting error in older adults, with these effects being greater in the trained limb than in the untrained limb, at low (20% MVC) rather than high (40% MVC) forces and in sinusoidal rather than constant force production conditions. Alterations in upper-limb strength, as well as the power spectral, regularity, force share, and/or coupling of the digit forces, were expected to be associated with the training group’s enhanced force control.

METHODS

Participants

We recruited 14 subjects for this 6-week study. Determination of the number of subjects needed was based on a power analysis using data from Bilodeau et al’s study of subjects who completed a 4-week strength-training program. This analysis revealed that 8 subjects in each group (strength training, control) were required to attain 80% power with a P value of .05. Entry criteria included being male, aged between 70 and 80 years, physically active and in apparent good health, not involved in a strength-training program in the last year, with normal vision or vision corrected with eye glasses, and with no self-reported neurologic, musculoskeletal, sensory, or cognitive impairments that could affect training or testing performance.

The first 7 subjects who gave informed consent were assigned to the strength-training group (age, 75±2y; height, 171±6cm; mass, 72±7kg; body mass index [BMI], 25±3kg/m²); the next 7 subjects were assigned to the control group. Three control subjects dropped out after the pretest sessions; consequently only 4 control subjects (age, 76±2y; height, 180±5cm; mass, 91±9kg; BMI, 28±4kg/m²) completed both testing sessions. The Griffith University Human Research Ethics Committee approved the study.

Experimental Design

We assessed the effect of a 6-week unilateral, dumbbell-based strength-training program on the isometric tri-digit finger-pinch force control of older men. Changes in force output were assessed in 4 finger-pinch force conditions that represented all combinations of 2 independent measures—mean force level (20% MVC, 40% MVC) and target shape (constant, sinusoidal). The sinusoidal target shape waveforms oscillated at 0.1Hz with an amplitude of ±5% MVC around the mean forces of 20% MVC and 40% MVC. Each of the 4 finger-pinch conditions (20% MVC, 40% MVC, 20%±5% MVC, 40%±5% MVC) were performed unilaterally, resulting in 8 force production conditions. Six 20-second trials were performed for each condition in both the pre- and post-training testing sessions, with each condition performed in randomized blocked order.

The goal of the finger-pinch conditions was to match the total finger-pinch force to that of the target force displayed on a computer monitor. Individual digit force outputs (thumb, index finger, middle finger), as well as the surface electromyographic activity of the flexor pollicis brevis (FPB) and flexor digitorum superficialis (FDS) muscles, were recorded during all trials.

Strength Training Procedures

Subjects in the strength-training group participated in 2 training sessions a week, with each session separated by 72 to 96 hours. The training program consisted of 3 unilateral dumbbell exercises (bicep curls, wrist flexions, wrist extensions), with the arm selected for training randomly determined for each subject. Three subjects trained with their preferred limb and 4 with their nonpreferred limb.

After a 5-minute arm-cranking warm-up on a modified Model 818E Monark cycle ergometer, subjects completed 4 sets of each of the 3 exercises. The first set was used as a specific warm-up, with a load of about 40% to 50% of the 5-repetition maximum (5-RM). As recommended by Fiatarone-Singh, the remaining 3 sets of each exercise were performed with loads that could only be lifted for 8 to 10 repetitions per set. Once the subjects could perform 4 sets of 10 repetitions each, one of the tasks was increased by 1kg at the next training session. Subjects rested for 2 to 3 minutes between sets. Control subjects did no training and maintained their normal activities during the study period.

Assessment Procedures

Finger-pinch force control. When performing the tri-digit finger-pinch force production tasks, subjects were seated about 80cm from a 43-cm (17-in) computer monitor positioned at eye level. The monitor showed the target force in full before each trial began, with the total finger-pinch force appearing in real-time superimposed over the target force. During testing, the selected upper arm was flexed to 90° at the elbow, with the forearm supported on an armrest and held in a neutral (pronation, supination) position. The subject grasped the load cell transducers (total mass, ~250g) with the thumb, index finger, and middle finger. In this position, the pulp of the thumb was positioned over the center of one of the scale markers. The index and middle fingers were positioned on 2 BC302 117.6 N load cell transducers, respectively. In this arrangement, the sum of the 2 force forces theoretically equaled that of the total (thumb) force. Kinetic data were sampled at 1000Hz and filtered using a second-order Butterworth low-pass filter (cutoff frequency, 20Hz). Before each testing session, the 3 load cells were calibrated using a 2-point procedure whereby the voltage produced from 2 known loads was recorded. After calibration, 3 unilateral finger-pinch MVCs were performed for each limb to determine the 20% and 40% MVC loads.

Surface muscle electromyographic activity of the FPB and FDS muscles was recorded using MediTrace bipolar surface Ag-AgCl electromyography electrodes. The electrodes were placed on the belly of the FPB and FDS muscles (interelectrode distance, 1.5cm) in parallel with the orientation of the underlying muscle fibers. All electromyographic signals were sampled at 1000Hz, amplified with V75-02 Coulbourn isolated bioamplifiers, and transferred to a computer for further analysis. Electromyographic data were subsequently filtered with a second-order Butterworth low-pass filter with the cutoff frequency set at 400Hz.

Upper-limb strength. Isometric tri-digit finger pinch strength was determined via an MVC. Dynamic upper-limb strength was assessed using the 3 training exercises (dumbbell bicep curl, wrist flexion, wrist extension) with a 5-RM protocol similar to that described by Hrysomallis and Kidgell. The finger-pinch MVCs were performed at the beginning of each testing session and followed the same procedures used in the assessment of finger-pinch force control. After performing several submaximal isometric finger-pinch contractions, each subject performed 3 MVC trials with each limb. Each contraction was held for about 4 seconds, with a 60-second rest between each trial. The highest force value recorded among the 3 trials for a particular limb was considered the limb-specific MVC.

Dynamic (5-RM) upper-limb strength was assessed after finger-pinch force control. During all 5-RM assessments, sub-
jicts were seated with the upper body fully supported. All subjects completed a 5-minute arm-cranking warm-up on the Monark ergometer at 30W, after which they performed 2 warm-up sets of progressively heavier bicep curls, each set comprising 5 repetitions. The 5-RM load was defined as the heaviest mass successfully lifted for 5 repetitions. After the 5-RM bicep curl strength of the preferred and nonpreferred limbs was determined, this process was repeated for the wrist flexion and wrist extension exercises. Ignoring the 2 warm-up sets, 2 to 3 attempts were generally required to obtain the 5-RM for each exercise.

Data Analysis

Kinetic and electromyographic data collection began when the tri-digit finger-pinch force output exceeded 95% of the required (target) force. All data analyses were performed using custom-written Matlab software.11

**Force and electromyographic measures.** We assessed finger-pinch force control by calculating the level of relative force variability and targeting error.11,13 Because the effect of strength training on finger-pinch force control was assessed in both constant and sinusoidal force production conditions, the total force output was detrended before analysis so that any variability in force output would reflect that of the total, not the target force. We then calculated the force variability as the standard deviation of the detrended finger-pinch force-time-series divided by the mean target force. Targeting error was equal to the mean absolute difference between the total and target force (root mean square [RMS] error) divided by the mean target force. The amplitude of the electromyographic signals was obtained by using the RMS procedure with a bin size of 100ms.

We performed a power spectral analysis on the total pinch force between 0 to 4Hz because most (>90%) power during digit force production is found within this bandwidth.7,11,31 Within this frequency range, peak power frequency and proportional power were calculated. We examined the regularity of the detrended total finger-pinch force by calculating the sample entropy.32 Sample entropy is analogous to the more commonly performed approximate entropy,33-35 but is less affected by record length and has greater relative consistency.32 Sample entropy is defined as the natural logarithm of the conditional probability that 2 time series that are similar across a certain distance remain similar across the next distance.32 Sample entropy analysis outputs a value between 0 and 2, where lower values indicate greater signal regularity.

**Coupling relations.** Coupling was determined using inter-digit force sharing and cross-correlation analyses. Force sharing was calculated by expressing the force output of the subject fingers (index and middle) as a percentage of total finger-pinch force.3 We used cross-correlation analysis to quantify the

![Sample](image-url)
strength of the coupling between: (1) the target and 3 digit forces (target-digit force coupling); (2) the 3 digit forces (interdigit force coupling); and (3) the 2 RMS electromyographic (FPB, FDS) signals and 3 digit forces (electromyographic activity digit force coupling). The peak cross-correlation within a ±5-second time lag was recorded for all coupling measures.

Statistical Analysis

We used a 1-way analysis of variance (ANOVA) to compare the baseline results for all subjects. Because there were several significant intergroup differences at baseline, we performed repeated-measures analyses of covariance (ANCOVAs) to assess the effect on the dependent variables of 1 between factor (strength training) and up to 3 within factors (limb, mean force level, target shape). The effect of mean force level and target shape was determined by grouping the appropriate conditions (20% MVC, 40% MVC, 20% ± 5% MVC, 40% ± 5% MVC). The between-group difference (strength training vs control group) for the change (post-test minus pretest) scores was used as the effect statistic in all ANCOVAs. The between-group differences were calculated separately for the trained and untrained limb of the strength-training group. In these comparisons, the data for both the preferred and nonpreferred limbs of the control subjects were collapsed (averaged) so that the between-group comparisons involved comparing the unilateral change scores for the strength-training group with the control group’s bilateral (preferred and nonpreferred limb) mean change scores. Fixed effects in the ANCOVA model were all possible main effects and interactions between group (control, strength-training), limb (trained, untrained), mean force (20% MVC, 40% MVC), and target shape (constant, sinusoidal). Pretest scores were used as a covariate in all ANCOVAs to control for any between-group differences at baseline.

All dependent variables were log-transformed before being entered into the ANCOVA to reduce nonuniformity and to express the between-group change scores as percent differences.
The effect of strength training on the dependent variables was determined by calculating Cohen effect sizes. Effect sizes of less than .20, less than .50, less than .80, and greater than .80 represented trivial, small, moderate, and large effects, respectively. The between-group change score data are reported as the mean ± 95% confidence limit and are reported as effect sizes in the text. Data shown in the figures are expressed as percent change scores. All analyses were performed using SAS statistical software, with significance set at \( P \) less than .05.

**RESULTS**

The attendance rate for the 7 subjects in the strength-training group was 98.8%. Results of the baseline (pretraining) between-group ANOVA revealed that the strength-training group had significantly less body mass and upper-limb strength than the control group (\( P \) range, .001-.042). Although it was not statistically significant (\( P \) range, .193-.197), the strength-training group also tended to have greater baseline levels of force variability and targeting error. Table 1 shows the pre- and post-test scores for the finger-pinch force characteristics and upper-limb strength tests.

**Upper-Limb Strength**

We found no significant change in the control group in bicep curl, wrist flexion, wrist extension, or finger-pinch MVC strength. In comparison, the strength-training group experienced significantly greater increases in the trained limb’s bicep curl (effect size, 1.31±1.30; \( P = .049 \)) and wrist flexion (effect size, 2.85±1.49; \( P < .001 \)) strength, as well as in the wrist flexion strength of the untrained limb (effect size, 2.81±1.50; \( P < .001 \)). While there were moderate-to-large increases in the trained limb’s wrist extension (effect size, 1.00±1.07; \( P = .065 \)) and finger-pinch MVC strength (effect size, 0.70±0.75; \( P = .057 \)) for the strength-training group compared with control group, these differences were not statistically significant. Figure 1 shows the strength percent change scores for the trained limb of the strength-training group compared with the bilateral
(preferred and nonpreferred limb) mean values for the control group.

**Force Characteristics**

Figure 2 shows the representative pretraining time- and frequency-domain finger-pinch force data for the 2 groups performing constant and sinusoidal finger-pinch force production conditions.

Finger-pinch force variability and targeting error did not change significantly in the control group. The trained limbs in the strength-training group had a significantly greater reduction in force variability (effect size, $-0.27 \pm 0.17$; $P = .003$) and targeting error (effect size, $-0.39 \pm 0.19$; $P < .001$). There were no significant differences, however, for the untrained limb between the 2 groups with regard to the change in force variability (effect size, $-0.11 \pm 0.33$; $P = .487$) or targeting error (effect size, $-0.15 \pm 0.24$; $P = .216$). Subsequent analysis revealed that the reduction in force variability for the trained limb in the strength-training group was only statistically significant at high forces (40% MVC) (effect size, $-0.39 \pm 0.31$; $P = .017$) and during sinusoidal force production (effect size, $-0.38 \pm 0.31$; $P = .019$). The reduction in the trained limb’s targeting error was statistically significant at both low and high forces as well as during constant and sinusoidal force production (effect size, $-0.26 \pm 0.21$ to $-0.52 \pm 0.31$; $P$ range, .002–.014).

There were no significant changes in any of the finger-pinch force power spectral parameters (peak power frequency or proportional power) or sample entropy measures in the control group. Also, there were no significant between-group differences in the change scores of the power spectral parameters (effect size, $-0.42 \pm 0.61$ to $-0.39 \pm 0.58$; $P$ range, .151–.952).

In the strength-training group, force sample entropy increased to a significantly greater extent compared with the control group for both the trained (effect size, $0.78 \pm 0.50$; $P = .003$) and untrained limb (effect size, $0.62 \pm 0.50$; $P = .024$). Subsequent analysis revealed the increase in the trained limb’s sample entropy values to be significant at high forces (effect size, $1.08 \pm 0.85$; $P = .015$) and during sinusoidal force production (effect size, $0.93 \pm 0.58$; $P = .003$). Figure 3 shows the force variability, targeting error, and sample entropy percent change scores for the trained limb of the strength-training group compared with the bilateral (preferred and nonpreferred limb) mean of the control group.

**Coupling Relations**

The control group had significant decreases in the force share of the index finger and significant increases in the force share of the middle finger between the pre- and post-testing sessions. Consequently, the decrease in the index force finger share and the increase in middle finger force share in the strength-training group were significantly less than that of the control group for both the trained (effect size, $-0.77 \pm 0.65$; $P = .022$) and untrained (effect size, $-1.00 \pm 0.56$; $P = .001$) limbs. Further analyses revealed the force share change scores of the strength-training group’s trained limb were significantly less than the control group’s change scores at high forces and in constant and sinusoidal force production conditions (effect size, $-0.99 \pm 0.96$ to $-1.31 \pm 0.73$; $P$ range, .001–.044). The strength of many of the target-digit, interdigit, and electromyographic activity digit force coupling measures, as assessed by cross-correlation analysis, did not change appreciably for either group. As a result, there were relatively few significant between-group differences in the digit force coupling measures. Table 2 shows the pre- and post-test scores for interdigit force sharing and force coupling (target-digit, interdigit, electromyographic activity digit force) for both groups.

The strength-training group had significantly greater increases in the trained limb’s target-index finger (effect size, $0.11 \pm 0.08$; $P = .008$) and FPB-index finger (effect size, $0.57 \pm 0.54$; $P = .039$) coupling than did the control group. In contrast, the change scores for the strength-training group’s untrained limb were significantly smaller than those of the control group for target-middle finger (effect size, $-0.59 \pm 0.35$; $P = .002$), thumb-middle finger (effect size, $-1.05 \pm 0.77$; $P = .009$), and FDS-middle finger (effect size, $-0.48 \pm 0.36$; $P = .011$) coupling.

**DISCUSSION**

We conducted this study primarily to determine whether a general (nonspecific), unilateral, dumbbell-based strength-training program would significantly improve the isometric finger-pinch force control (force variability, targeting error) of older men. Changes in digit force output were assessed across 8 finger-pinch conditions reflecting all combinations of limb (trained, untrained), mean force level (20% MVC, 40% MVC), and target shape (constant, sinusoidal). We hypothesized that strength training would significantly improve finger-pinch force control and that this improvement would be associated with increased upper-limb strength as well as with changes in the power spectral, regularity, force share, and/or coupling of the digit forces.
Results have been reported in previous strength and coordination-training studies. Those 3 studies all used testing (transfer) tasks that were similar to the exercises performed in training. In contrast, we used a more commonly performed but nonspecific upper-limb strength training protocol and assessed the resultant changes in fine-motor control (isometric tri-digit finger-pinch force production). Our results show that training-related improvements in older adults’ force control resulted as a consequence of performing functional, multiple finger tasks that were not necessarily practiced during training. This generality of response may have implications for the design of exercise training programs for this population because older adults may not have sufficient time to practice every ADL in which an age-related loss of digit force control compromises their performance.

Our secondary aim in this study was to determine if unilateral strength training could produce any cross-education effects, that is, improve the finger-pinch force control and upper-limb strength of the untrained limb. There was no significant reduction in the untrained limb’s finger-pinch force variability or targeting error. There was a significant increase in the untrained limb’s wrist flexion strength, along with moderate (effect size >0.5) although nonsignificant increases in the untrained limb’s bicep curl and wrist extension strength. Consistent with previous research, our results indicate that older people retain sufficient neural plasticity to gain some significant cross-education strength effects from unilateral strength training. Further research is required to determine whether such cross-education effects can also be achieved in both fine- and gross-motor tasks.

### Effect of Force Level and Target Shape

The reduction in the strength-training group’s trained limb’s force variability was statistically significant at high (40% MVC), but not at low (20% MVC), forces and during sinusoidal but not during constant force production. In contrast, the reduction in targeting error was significant at both low and high forces and in constant and sinusoidal force production conditions. This difference in the task-dependency of the training response of force variability to targeting error may reflect differences in the manner in which these variables are calculated or controlled. It is also possible that the significant reductions in targeting error across all levels of force and target shape reflected the greater potential for change in targeting error than in force variability. This can be seen in the baseline data of our study and that of Keogh et al., in which there were greater between-group and age-related differences, respectively, in targeting error than in force variability. Because the age-related increase in force variability is more pronounced during sinusoidal force production than during constant force production, the greater reduction in force variability during sinusoidal than during constant force production tasks could also reflect older adults’ greater potential for improvement in sinusoidal force control. The greater reduction in force variability at higher forces was unexpected because previous studies found only training-related reductions in force variability at forces of 20% MVC or less. Such a discrepancy may reflect intersubject differences in the range of force outputs.
tested, as well as the exercises performed in training. Bilodeau and Ranganathan and colleagues only assessed changes in force variability at low forces (≤20% MVC), therefore the effect of training on higher-level (>20% MVC) force control remains unclear. The training exercises used in these studies, that is, index-finger abduction and dexterity ball manipulation, are both fine-motor movements that require very low absolute force outputs. In contrast, the training exercises we used could be considered more gross-motor movements because they involve multiple, larger muscles and produce considerably greater overall forces. Our results are therefore actually consistent with the specificity principle, because the changes in performance (in this case force variability) were more pronounced at high than at low forces, that is, at forces more similar to those performed in training.

Mechanisms Underlying Improved Force Control

While many mechanisms have been proposed as contributors to the age-related decline in digit force control, Sosnoff and Newell recently postulated that the primary mechanism contributing to this loss of force control is an age-related decline in muscular strength. The current and previous studies have shown that improvements in older adults’ force control have occurred in association with increases in upper-limb strength. Inasmuch as greater improvements in force control may occur in older adults with lower initial strength and/or force control, merely increasing upper-limb strength has the potential to lead to improvements in their digit force control capabilities.

Consistent with our initial hypothesis, the strength-training group also had a significantly greater increase in force sample entropy (irregularity) than did the control group. A highly regular signal (indicated by a low sample entropy or approximate entropy value) is thought to characterize an output produced from a small number of components or processes. It has been proposed that increases in force irregularity may contribute to the improvements in force control by increasing the number of force production modes available and hence, the adaptability and flexibility in precisely controlling the digit forces.

The lack of a significant between-group difference in the peak power frequency or proportional power change scores observed in this study was unexpected because increases in force irregularity are typically associated with an increase in peak power frequency and/or decreased low-frequency proportional power. Consistent with our findings, however, Newell et al found significant practice-related changes in targeting error and digit-force irregularity, with no significant change in the peak power frequency.

The between-group differences in the change scores for the digit force sharing and coupling variables were also not consistent with our initial hypotheses. As expected, the strength-training group had significantly greater changes in 2 digit-force coupling measures for the trained limb than the control group. This differed from the results for the index and middle finger (both trained and untrained limbs) force share and 3 digit-force coupling measures (untrained limb only), where the change scores for the strength-training group were significantly less than those of the control group. Notwithstanding the limitations of the reduced size of the control group, the relative lack of change in these measures for the strength-training group may reflect the short duration and/or the generality (nonspecific nature) of the training program. Whereas the training program involved dynamic bicep curl, wrist flexion, and wrist extension exercises, we assessed force control using isometric finger-pinch contractions. Alterations in the finger-pinch force share and/or coupling patterns may therefore only occur as a consequence of a more specific training program, that is, one involving isometric finger-pinchings.

Study Limitations

A potential limitation to the widespread application of our findings, however, is the between-group differences in strength and force control at baseline. At baseline, the strength-training group had greater levels of force variability and targeting error (with this difference still somewhat apparent post-training) and less upper-limb strength than the control group. The strength-training group at baseline also had more force variability and targeting error and less finger-pinch strength than the age-matched normative data reported by Keogh et al. Thus, consistent with the findings of Bilodeau et al and Sosnoff and Newell, resistance-training-related improvements in digit force control may be more pronounced in older adults who have reduced levels of upper-limb force control and/or strength.

CONCLUSIONS

Six weeks of a unilateral dumbbell-based strength training program produced a small but statistically significant reduction in finger-pinch force variability and targeting error in the trained limbs of a group of older men who had elevated baseline levels of force variability and targeting error. Because the training exercises differed from the testing (transfer) tasks, these results show that improvements in fine-motor control can be gained through the performance of more global, gross-motor tasks involving moderately large muscle groups. Our results further support the prescription of strength training to improve movement function in older adults, particularly in those with reduced baseline function.

Acknowledgments: We thank Will Hopkins, PhD, for his invaluable contribution to the development and implementation of the statistical analysis model used in this study. Thanks also to Justin Kavanagh, PhD, and Patria Hume, PhD, for comments on earlier drafts of the study.

References

15. Laidlaw DH, Bilodeau M, Enoka RM. Steadiness is reduced and motor unit discharge is more variable in old adults. Muscle Nerve 2000;23:600-12.

Suppliers
b. Applied Measurement, 14 Dalgety St, Oakleigh, Vic 3166, Australia.
c. Australasia Scales, 1-944 Nudgee Rd, Northgate, Qld 4013, Australia.
d. Medeleq, 11 Palings Ct, Nerang, Qld 4211, Australia.
e. Coulbourn, 7462 Penn Dr, Allentown, PA 18106.
f. Release 12, version 6.0; The MathWorks, 3 Apple Hill Dr, Natick, MA 01760-2098.
g. SAS Institute, 300 Burns Bay Rd, Private Bag No 52, Lane Cove, NSW 2066, Australia.
What Do We Really Know About the Transition to Adult-Centered Health Care? A Focus on Cerebral Palsy and Spina Bifida

Jessie A. Binks, MSc, Wendy S. Barden, MSc, Tricia A. Burke, BA, Nancy L. Young, PhD


Objectives: To address the lack of synthesis regarding the factors, processes, and outcomes specific to the transition from child-centered to adult-centered health care for people with cerebral palsy (CP) and spina bifida (SB); more specifically, to identify barriers, to outline key elements, to review empirical studies, and to make clinical and research recommendations.

Data Sources: We searched Medline and CINAHL databases from 1990 to 2006 using the key words: transition, health care transition, pediatric health care, adult health care, health care access, health care use, chronic illness, special health care needs, and physical disability. The resulting studies were reviewed with a specific focus on clinical transition for persons with CP and SB, and were supplemented with key information from other diagnostic groups.

Study Selection: All studies meeting the inclusion criteria were included.

Data Extraction: Each article classified according to 5 criteria: methodology, diagnostic group, country of study, age group, and sample size.

Data Synthesis: We identified 149 articles: 54 discussion, 21 case series, 28 database or register, 25 qualitative, and 34 survey articles (some included multiple methods). We identified 5 key elements that support a positive transition to adult-centered health care: preparation, flexible timing, care coordination, transition clinic visits, and interested adult-centered health care providers. There was, however, limited empirical evidence to support the impact of these elements.

Conclusions: This review summarizes key factors that must be considered to support this critical clinical transition and sets the foundation for future research. It is time to apply prospective study designs to evaluate transition interventions and determine long-term health outcomes.

Key Words: Cerebral palsy; Disabled children; Health care systems; Rehabilitation; Spinal dysraphism.

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

There have been impressive increases in the life expectancy of children with cerebral palsy (CP) and spina bifida (SB) in the last 2 decades as a result of improvements in clinical care.1-3 These gains in life expectancy present new challenges, including the transition to adult-centered health care,4 that often occurs between 16 and 21 years of age, when the subjects become ineligible for children’s services. Adults must move beyond child-centered health care to preserve these limited resources for children,7-11 and to ensure that adults receive age-appropriate services such as routine health promotion and screening.

The challenge of this transition, for youths with a variety of chronic health conditions, received international attention in the United States at the 1989 Surgeon General’s conference on “Growing up and Getting Medical Care: Youth with Special Health Care Needs.”12 Other groups have also recognized this challenge.9,13 Because of this attention, transition programs are being developed to bridge the gap between child and adult-centered health care.1,5,9,14-16 These programs are meant to give youths the knowledge and skills required to negotiate the adult health care system independently,17-21 to minimize or prevent secondary illnesses,19,22 to promote autonomy, and facilitate their maximum potential.12,23-24 They are also intended to plan for long-term needs, to give youths hope and focus,12,25,26 and to teach them that leaving pediatric care is not a precursor to drastic declines in health.27-28

The transition to adult-centered health care is uniquely challenging for people with CP24,29,30 and SB31,32 because their complex health care needs become superimposed on an adult health care system with fragmented services.4 The adult system is often lacking in expertise specific to chronic conditions of childhood, and rarely includes multidisciplinary teams that are central to the care of people with CP and SB.18,31,33,36 Because the health and well-being of these vulnerable adults hinges on uninterrupted access to health care,37 a successful clinical transition is critical.38

Although some key elements essential to transition services have been proposed in the literature, a synthesis of this information as it relates to CP and SB is not available to guide clinicians in the development of evidence-based programs. Furthermore, few standardized programs have been implemented and empirical studies have rarely evaluated the effectiveness of such programs.26,39 The purpose of this comprehensive review was to synthesize the key issues, to facilitate the development of clinical programs, and to direct future empirical research specific to CP and SB.

Our objectives in this review were (1) to identify barriers associated with the transition to adult-centered health care for CP and SB patients; (2) to outline key elements that should be included in transition programs for these groups; and (3) to...
METHODS

We conducted a comprehensive review of the clinical literature, using Medline and CINAHL databases, encompassing the period 1990 to 2006. We limited our review to English-language studies published in the peer-reviewed literature and we applied the following 4 search strategies.

First, in our database search we used the following key words: transition, health care transition, pediatric health care, adult health care, health care access, health care use, chronic illness, special health care needs, and physical disability. The results were combined (using the and command) with the terms cerebral palsy or spina bifida.

Second, we used an author name search of the same databases to find additional studies by authors identified through the primary search strategy. These articles were only included if they were focused on the transition to adult-centered health care for people with CP and SB.

Third, the reference lists from all articles found resulting from the first 2 strategies were reviewed. We did this to identify key articles not found in the original search.

Fourth, studies related to health care transitions of people with other conditions previously experienced only in childhood, but now common in adulthood, were also identified. These included: diabetes, respiratory disease, renal impairment, juvenile idiopathic arthritis (JIA), inflammatory bowel disease (IBD), cystic fibrosis (CF), congenital heart disease (CHD), celiac disease, cancer, organ transplant, and epilepsy. These articles were included to augment the limited literature with a focus on CP and SB. Studies focused on conditions other than CP or SB (ie, articles identified by strategies 2, 3, and 4), were included in this review only if they provided key information related to clinical transition that was considered relevant to CP or SB.

All studies that met the inclusion criteria specified above were reviewed jointly by 2 reviewers and labeled according to the reviewers’ consensus on 5 criteria that described the nature of each study: methodology, diagnostic group, country of study, age group, and sample size. The labels were generated based on the themes that were common in the articles reviewed. Some studies included multiple methods, diagnoses, or age groups. In such cases, the study was included in all relevant groupings. This information was compiled in table format to describe the sample of articles used in this review.

The main review process included all articles, which were coded according to their main messages. Articles with similar main messages were placed together in common themes. Again, many articles contributed to more than 1 theme. Article themes were organized under the main objectives and identified a priori: barriers, key program elements, and empirical evidence. Details regarding the main themes were synthesized and are presented in Table 1.

RESULTS

Our review of the literature identified 149 articles that were published between 1990 and that were relevant to the transition to adult-centered health care. (The reference list includes 150 references because of the inclusion of reference 12, which is a conference proceeding that was a keynote in the transition literature, but that was not peer-reviewed. Peer review was a criterion for inclusion in the results.) Of the 149 articles, 36% were discussion studies, 17% reported the results of qualitative interviews, 23% reported survey results, 19% reported on secondary analyses of pre-existing databases or registers, and 14% were case series reports. Most of the studies were conducted in the United States, the bulk of the remainder were done in the United Kingdom, Canada, and Australia. Table 2 summarizes the articles included in this review.

Several categories appeared consistently across the articles. These were organized under each of the objectives set a priori. The common themes and the number of articles that contributed to each theme are presented in Table 1.

Objective 1: Barriers to Transition

Many barriers to transition for people with CP and SB were identified; most were similar to those experienced by adults with other complex chronic illnesses. Barriers were reported from several perspectives, including providers of child-centered health care, youths in transition, parents of these youths, and providers of adult-centered health care.

Child-centered health care providers. One of the greatest barriers to an effective transition was reported to be the inability of child-centered health care professionals (eg, pediatricians) to “let go” of their long-standing relationships with patients,1,3,10,11,14,15,18,20,23,25,27,28,31,40,45 and their distrust of adult-centered health services.1,10,14,20,23,25,26,31,41,46 This often results in a lack of planning and a critical delay in readiness to transfer.20,45,46 The literature, however, underscores the importance of leaving the child-centered health care system because child-centered providers may not be best suited to care for young adults.10 For example, pediatricians often find it difficult to discuss adult issues,26,40,47 and may be inexperienced in dealing with changes in childhood diseases during adulthood.15

Table 1: Number of Articles per Theme

<table>
<thead>
<tr>
<th>Theme</th>
<th>No. of Articles Cited</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child-centered health care providers</td>
<td>22</td>
</tr>
<tr>
<td>Youth</td>
<td>16</td>
</tr>
<tr>
<td>Parents</td>
<td>23</td>
</tr>
<tr>
<td>Adult-centered health care providers</td>
<td>18</td>
</tr>
<tr>
<td>Total no. of articles contributing to objective 1 (barriers to transition)</td>
<td>30*</td>
</tr>
<tr>
<td>Timing</td>
<td>19</td>
</tr>
<tr>
<td>Preparation</td>
<td>33</td>
</tr>
<tr>
<td>Coordinated approach</td>
<td>33</td>
</tr>
<tr>
<td>Transition clinics</td>
<td>27</td>
</tr>
<tr>
<td>Interested adult-centered health care providers</td>
<td>23</td>
</tr>
<tr>
<td>Total no. of articles contributing to objective 2 (key elements suggested for transition programs)</td>
<td>42*</td>
</tr>
<tr>
<td>Introduction</td>
<td>56</td>
</tr>
<tr>
<td>Empirical evidence based on other chronic conditions of childhood</td>
<td>41</td>
</tr>
<tr>
<td>Why does evidence from other populations not apply to CP and SB?</td>
<td>42</td>
</tr>
<tr>
<td>Total no. of articles contributing to objective 3 (empirical evidence re: process and outcomes of transition)</td>
<td>123*</td>
</tr>
</tbody>
</table>

NOTE: Only 11 (7%) articles were original research specific to persons with CP and SB.
*The total number of articles contributing to each objective is less than the sum of the parts because some articles contributed to more than 1 part and were counted only once in the total.
Table 2: Summary of Articles Used in the Literature Review, With Emphasis on CP and SB

<table>
<thead>
<tr>
<th>Descriptor</th>
<th>Discussion Articles (n=54)</th>
<th>Database Analyses (n=28)</th>
<th>Case Series* (n=21)</th>
<th>Quantitative Methods* (n=34)</th>
<th>Qualitative Methods* (n=25)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Country</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>United States</td>
<td>34</td>
<td>United States</td>
<td>United States</td>
<td>United States</td>
<td>United States</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>13</td>
<td>United Kingdom</td>
<td>United Kingdom</td>
<td>United Kingdom</td>
<td>United Kingdom</td>
</tr>
<tr>
<td>Canada</td>
<td>4</td>
<td>Canada</td>
<td>Canada</td>
<td>Canada</td>
<td>Canada</td>
</tr>
<tr>
<td>Australia</td>
<td>3</td>
<td>Australia</td>
<td>Australia</td>
<td>Australia</td>
<td>Australia</td>
</tr>
<tr>
<td>Sweden</td>
<td>1</td>
<td>Sweden</td>
<td>Sweden</td>
<td>Sweden</td>
<td>Sweden</td>
</tr>
<tr>
<td>Italy</td>
<td>3</td>
<td>Italy</td>
<td>Italy</td>
<td>Italy</td>
<td>Italy</td>
</tr>
<tr>
<td>Germany/Netherlands</td>
<td>2</td>
<td>Germany/Netherlands</td>
<td>Germany/Netherlands</td>
<td>Germany/Netherlands</td>
<td>Germany/Netherlands</td>
</tr>
<tr>
<td>Japan</td>
<td>1</td>
<td>Japan</td>
<td>Japan</td>
<td>Japan</td>
<td>Japan</td>
</tr>
<tr>
<td>Norway</td>
<td>1</td>
<td>Norway</td>
<td>Norway</td>
<td>Norway</td>
<td>Norway</td>
</tr>
<tr>
<td>New Zealand</td>
<td>1</td>
<td>New Zealand</td>
<td>New Zealand</td>
<td>New Zealand</td>
<td>New Zealand</td>
</tr>
<tr>
<td><strong>Diagnostic group</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CP</td>
<td>6</td>
<td>CP</td>
<td>CP</td>
<td>CP</td>
<td>CP</td>
</tr>
<tr>
<td>SB</td>
<td>3</td>
<td>SB</td>
<td>SB</td>
<td>SB</td>
<td>SB</td>
</tr>
<tr>
<td>Ill-defined conditions</td>
<td>31</td>
<td>Ill-defined conditions</td>
<td>Ill-defined conditions</td>
<td>Ill-defined conditions</td>
<td>Ill-defined conditions</td>
</tr>
<tr>
<td>CF</td>
<td>6</td>
<td>CF</td>
<td>Diabetes</td>
<td>CF</td>
<td>Diabetes</td>
</tr>
<tr>
<td>JIA</td>
<td>1</td>
<td>JIA</td>
<td>Cancer</td>
<td>JIA</td>
<td>Cancer</td>
</tr>
<tr>
<td>Cancer</td>
<td>1</td>
<td>Cancer</td>
<td>CHD</td>
<td>Cancer</td>
<td>CHD</td>
</tr>
<tr>
<td>Other chronic conditions (IBD, organ transplant, renal and respiratory conditions)</td>
<td>5</td>
<td>Other chronic conditions (IBD, organ transplant, renal and respiratory conditions)</td>
<td>CHD</td>
<td>Other chronic conditions (IBD, organ transplant, renal and respiratory conditions)</td>
<td>CHD</td>
</tr>
<tr>
<td><strong>Age group</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Children</td>
<td>13</td>
<td>Children</td>
<td>Children</td>
<td>Children</td>
<td>Children</td>
</tr>
<tr>
<td>Youth and young adults</td>
<td>45</td>
<td>Youth and young adults</td>
<td>Youth and young adults</td>
<td>Youth and young adults</td>
<td>Youth and young adults</td>
</tr>
<tr>
<td>Adults</td>
<td>8</td>
<td>Adults</td>
<td>Adults</td>
<td>Adults</td>
<td>Adults</td>
</tr>
<tr>
<td>Sample size</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–30</td>
<td>NA</td>
<td>0–30</td>
<td>0–30</td>
<td>0–30</td>
<td>0–30</td>
</tr>
<tr>
<td>31–100</td>
<td>NA</td>
<td>31–100</td>
<td>31–100</td>
<td>31–100</td>
<td>31–100</td>
</tr>
<tr>
<td>201–300</td>
<td>NA</td>
<td>201–300</td>
<td>201–300</td>
<td>201–300</td>
<td>201–300</td>
</tr>
<tr>
<td>&gt;300</td>
<td>NA</td>
<td>&gt;300</td>
<td>&gt;300</td>
<td>&gt;300</td>
<td>&gt;300</td>
</tr>
</tbody>
</table>

NOTE. Items in bold indicate articles specific to CP and SB. Abbreviation: NA, not applicable.
*Some articles had methods and/or age ranges that overlapped, and therefore were included in all relevant categories.
Youth. Obstacles may also arise from youths themselves. There is often little incentive for them to move on from providers who have supported them very well for a long period. Young adults may be reluctant to leave the safety and familiarity of family-centered care. This philosophy is common practice in health care for children, but is rarely applied to adult-centered services. Some patients feel rejected or isolated after years of commitment by a multidisciplinary group of health professionals. Moving to adult services may also be perceived by some as a “step closer to disease complications and even death.”

Parents. Parents display their own distinct forms of resistance to their child’s clinical transition. Adult-centered services rarely engage with families in the same way as do child-centered services, and, as a result, parents may interfere with the transition because they feel excluded from the decision-making process. Although parents should be involved in planning, their inability to relinquish control over their child’s health and health care decisions may impede their child’s autonomy.

Adult-centered health care providers. According to the literature, adult-centered health care providers may present obstacles to successful transition. Adult-centered physicians may have limited training and experience with childhood chronic illnesses, and therefore have limited knowledge or interest in caring for these young adults. In addition, they may have insufficient resources to support the patient’s complex needs. For example, some patients noted in qualitative interviews that their physician was uncomfortable or ambivalent about some topics, had difficulty discussing bad news, and lacked specific training. Adult-centered providers also have a tendency to conduct a complete reassessment at the first appointment, which can be unsettling, exhausting, and frustrating for the young disabled person. The challenges of transition may be further complicated by poor communication between adult-centered providers and patients. Additionally, these clinicians often become overly focused on the patient’s chronic illness, and may fail to address the primary medical issues that prompted the visit, or fail to assess routine health and health promotion issues.

Objective 2: Key Elements Suggested for Transition Programs

Reducing or eliminating these transition barriers has been the impetus for a multitude of discussion articles suggesting essential components for successful transition programs. It is important to recognize that although these elements have been identified in the literature, most of the information is based on theory and has not been subjected to empirical evaluation.

Timing. The timing of the transfer should be established by both the pediatric provider and the family. Timing should be flexible, and should depend on the youth’s cognitive development, physical abilities, environment, and family support. Although no definitive time has been identified for actual transfer, a target age range is helpful in preparing for transition. Because many services may be unable to delay the age of transition, they may need to consider beginning the process earlier to align with an externally imposed transition deadline.

Preparation. There should be a long preparation period before leaving child-centered care. It is recommended that the transition process should be initiated years before the official transfer, and some suggest that it should begin at diagnosis. Patients must understand their condition and its daily management, and acquire basic skills to independently care for themselves. Nevertheless, independence does not mean “without assistance,” only that the young adult must take responsibility for the direction of his/her care, which may include directing an attendant to assist with specific activities. Continuous support and encouragement from family is also essential for the youth to attain autonomy and alter his/her self-perception from being “disabled” to being “competent.”

Coordinated approach. Transition planning for each patient should be a coordinated approach that includes feedback from the youth and his/her family, pediatrician, primary care physician, and adult specialist(s). The child-centered team should also prepare an up-to-date medical summary to be given to the adult-centered providers, including details of current treatment regimes and recommendations. Planning should also include the development of a clinical transition plan, and, at a minimum, include the services required and identify who will provide them. Youths with more severe functional limitations may require a plan developed over a longer period of time.

Financial planning (ie, estate details), education, and vocational training should also begin before adulthood.

Transition clinics. The transition process may be improved with several visits to a transition clinic before the actual transfer, and include a consult with both the child and adult-centered health care providers. Clinics can provide the patient and family with an overview of the general differences between the 2 sectors of care. They also serve as an avenue to provide ongoing support and resources to the family, such as contact with a mentor or peer group who have been through the process.

Young adults are encouraged to actively participate in decision making and to direct their own health care. Interested adult-centered health care providers. For all people involved, transition should be a gradual process, not an event, from child to adult-centered health care. There must be interest and collaboration between adult-centered health care providers on the receiving side of the transition. Adult-centered providers should not only care for the chronic condition, but also address basic issues such as independence, social interactions, body image, sexual health, alcohol and drugs, and anxiety. Transition planning must involve primary care physicians, which may provide the only medical continuity for young people and their families during a time of discontinuities. The child-centered team may need to find an adult-centered provider willing to accept the patient, and become a health care coordinator. Increasing ties between child and adult-centered departments may also promote confidence in the adult service. As the transition process unfolds, it should be monitored and evaluated to help guide the development of best practice.

The key limitation of many of these studies is that the elements have not been implemented consistently in transition programs, nor formally evaluated. Therefore, they are based on theory and clinical experience and are not evidence-based. The volume of discussion articles over the past decade, however, has stimulated an interest in evaluation and, as a result, some qualitative and quantitative studies on clinical transition are now being done.
Objective 3: Empirical Evidence Related to the Process and Outcomes of Transition

Very few empirical studies addressed the specific issue of clinical transition to adult-centered health care for CP and SB patients. Among these was a study by Stevenson et al. who assessed the use of health services, welfare, and social functioning before and after leaving school for youths and adults with CP. They identified fragmentation of services after adolescence. General health was considered poor in 21% of their “younger” group (20 and 22 years of age), and in 9% of their “younger” group (15 to 18 years of age). Their “younger” group also felt more socially isolated than the “younger” group. Morgan et al. assessed the decline in contact with health and social service departments for young adults with SB. It was clear from their evaluation that more than half of the young people had unmet medical needs and were grateful for the offer of an annual assessment in the adult setting.

Perceptions about clinical transition have been studied using qualitative methods. For example, Sawyer et al. described the planning and implementation of the transfer of 10 young people (mean age, 22y) with SB to an adult health care facility. They found that most participants had already transferred informally to an adult facility, but with little or no communication between the 2 systems. Darrah et al. found that families continued to experience dissatisfaction and frustration with service delivery. The consistent messages from these qualitative studies were that patients and caregivers were unprepared for their roles in the next health care setting, did not understand the essential steps in managing a chronic illness, and had limited access to appropriate health care practitioners for guidance. Effective clinical transitions must bridge the gap between health care institutions that often function in isolation. New models of “cooperative care” that link primary care providers and local services to regionalized adult-centered specialty services may make it possible to offer a meaningful transition experience to young people with chronic conditions.

Overall, there is limited empirical evidence related to the process and outcomes of the transition to adult-centered health care for CP and SB patients. Most of the empirical evidence relates to the functional status (ie, mobility) and social status (ie, living arrangements) of these populations, and their life expectancy and causes of death. A series of quantitative articles reported experiences of clinical transition, perceptions about the process from patient, parent, and physician perspectives, and identified issues associated with transition.

Empirical evidence based on other chronic conditions of childhood. Because of limited empirical evidence about transition for CP and SB patients, we elected to supplement the information with data related to transition from other populations with complex and chronic conditions of childhood. Empirical studies on the transition to adult-centered health care have most often focused on young people with CF, cancer, CHD, arthritis, and diabetes, where there are similar issues in adult populations and therefore have pre-existing adult specialists. A series of qualitative articles reported experiences of clinical transition, perceptions about the process from patient, parent, and physician perspectives, and identified issues associated with transition.

Examples from other populations that enrich our understanding of the clinical transition include a study conducted by Flume et al. in CF clinics across the United States. They assessed health care team members’ (ie, nurses’) perceptions of clinical transition. The authors stated that only 11% of team members agreed with an introduction to the idea of transfer at diagnosis, and many outlined valid criteria for transfer (ie, 27% cited pregnancy). Shaw et al. surveyed 17-year-old patients with JIA and their families, who had been transferred to adult care. They reported several ongoing transitional issues: 55% were still seeing the rheumatologist with a parent, 20% were not taking their medication independently, and 14% had not received career counseling.

Other studies used pre-existing databases to assess service use and the health status of populations with a chronic condition or special health care needs. For example, Lotstein et al. conducted a study to assess the proportion of youths with special health care needs who received services for medical transitions. Overall, 50% of parents had discussed their child’s changing needs with their child’s physicians, and 59% had developed a plan to address these needs. Other studies focused on transition to adult-centered health care, while others only addressed transition as an afterthought. Some studies evaluated follow-up procedures for patients who were discharged from a pediatric facility, and found that few programs focused on the long-term health needs of adults. Also, several qualitative studies assessed patient perspectives and experiences during their clinical transition, and their findings supported those of the quantitative papers cited above.

Why does evidence from other populations not apply to CP and SB patients? Although studies on other diagnoses provide useful information, they are limited in their generalizability to multidomain illnesses such as CP and SB. For example, CP or SB patients may present with limitations in communication, problem solving, learning, mobility, and feeding, and therefore require extensive health care support. The cumulative effect of these complex disorders makes transition difficult, and limits the patients’ ability to utilize the adult-centered health care system in the traditional patient-physician relationship.

The challenges of clinical transition in the CP and SB groups are also exacerbated by several factors, including the lack of multidisciplinary comprehensive health services for adults with a chronic illness. A dearth of adult providers with interest in chronic illnesses of childhood, and a lack of specialized training in the proper care of adults with CP and SB. Consequently, these adults are often left to navigate an unfamiliar and often ill-equipped health care system.

Many adults with CP and SB continue to attempt to access pediatric health care services despite their age and changing needs. Others struggle to connect to appropriate adult-centered health care resources. The remainder stop seeking medical attention, either out of frustration or rebellion. Clearly, there is a need for ongoing care because of ongoing health issues, coupled with the development of new health issues in adulthood.

DISCUSSION

We found sufficient discussion literature about clinical transitions. Currently, the need is to step beyond the concepts and into evidence-based research. This research must begin by defining the concept of success, as related to clinical transition, and how it can be effectively measured. Once this is achieved, suggested research questions include: (1) What is the distribution of health outcomes in adults with CP and SB? (2) What characteristics of transition are associated with better health in adults with CP and SB? (3) What patterns in health care utilization are associated with health and well-being of adults with CP and SB?

These questions may be best addressed by intervention studies that assess the population before and after clinical transition. Now is the time to measure health and well-being, when there are few standardized transition programs in place. Long-term studies will enable us to assess natural variation in the
population, and to tease apart patterns in health associated with the transition to adult-centered health care. People who are at the extremes of the health outcomes distributions, for example, the successes and failures, should be studied further to dissect the nuances of their transition process. Also, multiple methods should be combined within studies to generate a more comprehensive picture of the dynamics associated with transition. For example, qualitative methods enhance and add context to other data collection methods. This type of research will aid in the development of effective transition programs, inform policies on health care for these populations, and help secure funding for specialized services.

Future research must also pay attention to the full spectrum of their concept of the traditional clinician-patient relationship. Furthermore, clinicians must foster the autonomy of physically challenged young people in developing skills for clinician-patient communication. Clinicians should also create a complete discharge summary for both the patient and the adult-centered health care provider that serves as a health care “passport.” This will help empower young adults with knowledge about themselves and their disease, help them communicate with their new providers, and help reduce the time required during the reassessment phase. A basic health care passport may have the potential to address several key components of the transition to adult-centered health care described previously in this review.

Educational Implications for Health Care Providers

The key concepts regarding clinical transition, as well as the emerging empirical research on this topic, must also be translated into the curricula of nursing, rehabilitation, and medical schools. Programs should incorporate information on the unique challenges and special health care needs of CP and SB patients, including changes in the expression of childhood diseases with age. Curricula must also promote the development of skills in working with people who have cognitive and communication challenges, and foster an appreciation of the value and importance of having an advocate in attendance at clinical appointments. This may require clinicians to set aside their concept of the traditional clinician-patient relationship and to learn to include a third party. Seeing beyond the disability is critical in creating a culture of acceptance and instilling a sense of worth in these populations. Finally, providers must learn to balance the concurrent demands of chronic health care conditions with generic health promotion and screening needs.

Study Limitations

The primary limitation of this review is related to the source materials. The methods and data analysis sections of the source articles reviewed were often unclear (ie, may not have specified the measurement properties of their data collection tools). The different sampling strategies used in the source articles limited our ability to make comparisons across studies. We had difficulty applying standardized age groupings, because the source materials included variable age ranges and lacked details on the sample characteristics. The lack of detail regarding age within the source articles made it impossible to examine age-specific differences that are key to studying the impact of transition (see Table 1). The lack of detail on sample characteristics such as education, ethnicity, employment status, marital status, living arrangement, illness severity, and health status further limit cross-study comparisons.

CONCLUSIONS

The delivery of care to adults with conditions that have historically been limited to childhood presents new challenges to the health care system. There are many discussion studies on the transition to adult-centered health services that suggest key components for planning and designing successful clinical transition programs. Most studies, however, discuss key factors in isolation, without considering the multifaceted nature of several factors that have a concurrent impact. Furthermore, very few studies have applied and evaluated these concepts in practice, or assessed the direct impact of the transition to adult-centered health care on health outcomes. Moreover, empirical studies specific to CP and SB are rare and have been limited to a few health care systems, making it difficult to compare outcomes across systems.

Clearly, there is a clinical problem that must be addressed by the health care system to ensure that adults with CP and SB receive continuous, age-appropriate health care. The challenge now is to develop high quality, multidisciplinary care comparable to that available for children. A smooth and easy transition to adult-centered health care should be an expected and desired outcome of child-centered health care.

This review offers a significant step toward that goal by identifying key factors that must be considered in the planning of care for children and adults, and specifically for transition programs. It is time to use emerging theoretical models to develop transition interventions and to evaluate them using controlled trial methods that consider both medium and long-term health outcomes. This review shifts the focus to populations with conditions that require complex care and may have the most difficulty with clinical transition. The greatest opportunities for substantial improvement are found in these vulnerable populations.

Acknowledgments: We thank Aliza Sturm for providing support in the planning phase and Erika Schippel for her assistance in the editing phase of this review.

References


BRIEF REPORT

Energy Cost of the Case Western Reserve Standing Neuroprosthesis

George F. Forrest, MD, Thomas C. Smith, MD, Ronald J. Triolo, PhD, Jason P. Gagnon, PT, Darryl DiRisio, MD, Michael E. Miller, MS, Lori Murray, PT, John A. Davis, MD, Atif Iqbal, MD


Objective: To determine the oxygen consumption of a person with C7 American Spinal Injury Association (ASIA) grade B tetraplegia using the Case Western Reserve/Veterans Administration (CWRU/VA) standing neuroprosthesis.

Design: Measure the oxygen consumption and carbon dioxide production of a person with C7 ASIA grade B tetraplegia at rest, standing in the parallel bars with the CWRU/VA system on, ambulating in the parallel bars, and transferring from a wheelchair to a mat with the system on.

Setting: University medical center.

Participant: A 26-year-old man with C7 ASIA grade B tetraplegia. The subject was a recipient of the CWRU/VA standing neuroprosthesis.

Interventions: Not applicable.

Main Outcome Measures: Measurement of oxygen consumption and carbon dioxide production using a metabolic cart.

Results: Oxygen consumption of the subject was 1.22mL·kg⁻¹·min⁻¹ at rest. It was 4.7mL·kg⁻¹·min⁻¹ while standing in the parallel bars, 7.2mL·kg⁻¹·min⁻¹ while ambulating in the parallel bars, and 7.9mL·kg⁻¹·min⁻¹ when transferring from a wheelchair to a mat.

Conclusions: Oxygen consumption of the subject when using the system is about 2 metabolic equivalent units, which is compatible with sustained use of the system for standing.

Key Words: Oxygen consumption; Prostheses and implants; Rehabilitation; Spinal cord injuries.

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

From the Department of Physical Medicine and Rehabilitation (Forrest, Gagnon, Iqbal), Division of Pulmonary Medicine (Smith), and Department of Neurosurgery (DiRisio), Albany Medical College, Albany, NY; Functional Electric Stimulation Center, Consortium of Case Western Reserve University, Metro Health Medical Center and the Louis Stokes Cleveland Veterans Affairs Medical Center, Cleveland, OH (Triolo, Miller, Murray); and Department of Neurosurgery, Tulane Medical Center, New Orleans, LA (Davis).

Supported by the New York State Department of Health Spinal Cord Injury Research Board, the U.S. Food and Drug Administration Office of Orphan Product Development (grant no. FD-R-001244-08), Department of Veterans Affairs Rehabilitation and Development Merit Review (grant no. B3155R), and by the National Institute of Health (grant no. GCRC M01RR00800).

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.

Reprint requests to George P. Forrest, MD, Dept of Physical Medicine and Rehabilitation, Albany Medical College, 47 New Scotland Ave, MC 79, Albany, NY 12208, e-mail: forreg@mail.amc.edu.

0003-9993/07/8808-11074-6 $32.00/0

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

Arch Phys Med Rehabil Vol 88, August 2007

THE CASE WESTERN RESERVE/Veterans Administration (CWRU/VA) implanted standing neuroprosthesis is a device that allows people with low cervical or thoracic spinal cord injury (SCI) to exercise their legs, stand, and transfer.1-4 Some system recipients have been able to use the system to ambulate short distances with a walker and bilateral ankle-foot orthosis (AFO) by using a swing-to gait.5 The device consists of a radiofrequency transmitter (external control unit); a receiver stimulator that is implanted subcutaneously and attached to the fascia of the abdominal wall; a coupling coil that is taped onto the skin over the receiver and stimulator; and 8 electrodes that deliver stimulating current to motor points in the paraspinal muscles at L1, the gluteus maximus, the vastus lateralis, and the semimembranosus muscles bilaterally. The receiver stimulator is programmed by a laptop computer to provide each of the 8 muscles with the minimum amount of electric impulse necessary to elicit full contraction of the muscle. The amplitude of stimulation is 20mA. The frequency is 20Hz. The pulse width varies from 70 to 200µs as necessary to obtain maximum contraction of each muscle. The user activates the system by pushing a button on the external control unit or “finger-worn” command ring. The user can select programs to exercise the muscles individually or in groups. The program for standing elicits cocontraction of the 8 muscles resulting in extension of the hips, knees, and axial muscles. Standing, transfers, and ambulation require use of bilateral AFO and a walker. Waters and Mulroy,6 Hussey and Stauffer,7 and Solomonow et al8 have reported that the energy consumption required to use an orthosis is an important factor in determining whether a person with SCI will choose to use that device. The purpose of this study was to determine the energy used by a patient with tetraplegia by using the CWRU/VA neuroprosthesis. In particular, the investigators wanted to determine if the energy requirements of the system during transfers, standing, or gait would preclude functional use of the device.

METHODS

Patient Selection

The CWRU/VA neuroprosthesis can be used by patients with lower cervical or thoracic level SCI. Inclusion and exclusion criteria are listed in appendix 1. The patient in this study was a C7 American Spinal Injury Association (ASIA) grade B level injury.5 His injury occurred in 2002. His weight was 75kg. He had no medical problems other than SCI.

System Training

The usual protocol for use of the CWRU/VA neuroprosthesis requires the subject to use a surface stimulator to strengthen the target muscles for 8 weeks before the implantation of the system. The patient remains in the hospital for 3 to 7 days after the procedure. The patient is home with restricted activity from the time of discharge until 6 weeks after the time of surgery. At the end of week 6, the system is programmed so that each electrode provides the minimum amount of stimulus necessary
to elicit full contraction of its target muscle. During weeks 6 to 14, the subject uses the device at home every day to strengthen the muscles. The subject comes to clinic once a week. At the clinic, the investigator is able to check the function of each lead. The system records the time that it is used. This allows the investigator to verify compliance with the exercise protocols. At week 14, the subject begins physical therapy to learn to come from sit to stand, to transfer, and if possible to try to ambulate. The functional training period is listed in the protocol as week 14 to 26 but is continued as long as is necessary to allow the subject to maximize his/her use of the system. This subject had his device implanted on August 20, 2004. He began to use the system on September 30, 2004. The subject took a leave of absence from this research study on December 18, 2004, to participate in an unrelated clinical trial. He resumed participation in this study on February 1, 2005. He completed formal training in December 2005. On August 5, 2005, measurements were taken of the oxygen consumed by the subject using the CWRU/VA neuroprosthesis. A metabolic cart provided breath-by-breath analysis of oxygen consumed and carbon dioxide exhaled by the subject. The parameters measured included oxygen consumed in liters per minute (L/min) and oxygen consumed by weight (mL·kg⁻¹·min⁻¹). The measurements were taken with the patient sitting with the system off (2min, 30s), standing between the parallel bars (5min, 30s), ambulating between the parallel bars with a swing to gait (2min, 30s). The values reported are averages of the breath-by-breath analysis over the last 2 minutes of each stage.

RESULTS

The patient was able to stand for up to 2 hours. He was able to ambulate 8 to 15 feet in 1 minute (2.4–4.5m/min) (patient self-report, personal communication, May 9, 2006). The testing done on August 5, 2005, showed that the subject’s oxygen consumption while sitting with the system off was 1.22mL·kg⁻¹·min⁻¹. Oxygen consumption while standing between the parallel bars was 4.7mL·kg⁻¹·min⁻¹. Oxygen consumption while ambulating in the parallel bars was 7.2mL·kg⁻¹·min⁻¹. Oxygen consumption while transferring from a wheelchair to a mat was 7.9mL·kg⁻¹·min⁻¹.

DISCUSSION

The data indicate that the energy consumption of a man with tetraplegia using the CWRU/VA neuroprosthesis was not high. The energy consumption standing and transferring was 2 to 2.1 metabolic equivalents (METS), which is twice the energy use of a nondisabled person at rest. It is comparable to a nondisabled person ambulating at a rate of 2 miles an hour or playing the piano. The energy requirement is less than that of a paraplegic using a wheelchair propulsion with bilateral KAFO and less than that of a paraplegic patient ambulating with a hip-knee-ankle-foot orthosis (HKAFO). It is less than that of a reciprocating-gait orthosis with surface electric stimulation and similar to that of a person with paraplegia using the Parastep, a commercially available system of functional electric stimulation with surface stimulation. There are no prior reports of energy consumed by people with quadriplegia ambulating without body weight support by a frame. Carvalho et al reported that tetraplegic subjects using a system that included a harness for body weight support and surface stimulation over the quadriceps muscle and peroneal nerve had a mean oxygen uptake of 7.05mL·kg⁻¹·min⁻¹ to support a gait speed of 0.5km/h on a treadmill. Algood et al reported that the oxygen consumption of a tetraplegic using a manual wheelchair on a smooth level surface is 8.0mL·kg⁻¹·min⁻¹. Thus, the energy use of the CWRU/VA neuroprosthesis was quite comparable to that of use of a manual wheelchair (table 1).

The energy requirement of the patient in this study (2–2.1 METS) would not preclude the use of the system. Other major limitations could be local muscle fatigue and speed movement. During normal gait muscle, groups of the lower extremity alternately contract and relax through the gait cycle. The CWRU/VA neuroprosthesis requires cocontraction of the paraspinal muscles, hip extensors, and knee extensors to keep the user upright. The fact that this subject is able to stand for 2 hours would indicate that local muscle fatigue is not a barrier to use. The most efficient speed of walking for people without disability is about 75m/min. People with C7-8 tetraplegia have a speed of 50.6m/min in a standard-weight wheelchair and 55.4m/min in a lightweight wheelchair. The most comfortable ambulating speed for paraplegics using HKAFO with or without surface stimulation is 12m/min. The subject in this study reported a speed of ambulation 2.4 to 4.5m/min. This would make the system practical only for standing, transfers, and ambulation for very short distances.

Study Limitations

This study has a number of limitations. This is the first subject with quadriplegic tested for energy consumption standing, ambulating, and transferring with the CWRU/VA neuroprosthesis. The ability to use the system and the energy consumed by using the system may vary from subject to subject depending on upper-extremity and hand function, sensory sparing, balance, weight, and the presence or absence of other medical comorbidities. The testing was performed in accordance with study protocol that called for testing 1 year after the device was implanted. The patient had not yet fully completed his gait training or maximized his ability to use the walker.

The time of ambulation between the parallel bars may or may not have been long enough to conclude that the values of oxygen consumption measured represent steady-state values for use of the neuroprosthesis. A transfer from one surface to another is a short duration activity that does not lend itself well to measurement of steady-state energy consumption. The energy requirement for use of this neuroprosthesis as an aide to ambulation with a walker will be better tested by more sustained use of the device than the transfer that was performed in this study.

<table>
<thead>
<tr>
<th>Test</th>
<th>Oxygen Consumption (mL·kg⁻¹·min⁻¹)</th>
<th>METS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-disabled adult walking at comfortable rate</td>
<td>12.1</td>
<td>3.5</td>
</tr>
<tr>
<td>Non-disabled adult walking at fast rate</td>
<td>18.4</td>
<td>5.7</td>
</tr>
<tr>
<td>Paraplegic with bilateral KAFO</td>
<td>16.5</td>
<td>5.0</td>
</tr>
<tr>
<td>Paraplegic with bilateral HKAFO</td>
<td>9.61</td>
<td>2.7</td>
</tr>
<tr>
<td>Paraplegic with bilateral KAFO and surface stimulation</td>
<td>8–11</td>
<td>2.3–3.1</td>
</tr>
<tr>
<td>Parastep (standing)</td>
<td>7.85</td>
<td>2.2</td>
</tr>
<tr>
<td>Parastep (walking)</td>
<td>7.9</td>
<td>2.3</td>
</tr>
<tr>
<td>Tetraplegic gait with harness to support body weight</td>
<td>7.05</td>
<td>2.0</td>
</tr>
<tr>
<td>Tetraplegic wheelchair propulsion</td>
<td>8.0</td>
<td>2.3</td>
</tr>
<tr>
<td>CWRU/VA standing</td>
<td>4.2</td>
<td>1.2</td>
</tr>
<tr>
<td>CWRU/VA walking in parallel bars</td>
<td>7.2</td>
<td>2.1</td>
</tr>
<tr>
<td>CWRU/VA transferring wheelchair mat</td>
<td>7.9</td>
<td>2.3</td>
</tr>
</tbody>
</table>
CONCLUSIONS

The CWRU/VA neuroprosthesis allowed a patient with ASIA C7 grade B tetraplegia to stand and transfer. The energy consumption of the subject using the system would not prohibit the patient from standing for long periods of time and performing tasks such as reaching for a cup or a book with 1 hand. The subject can ambulate short distances, but the use of the system for ambulation is limited by the speed of movement, which is low compared with walking speeds of nondisabled people or the speed of a person with tetraplegia using a manual wheelchair.

APPENDIX 1: INCLUSION AND EXCLUSIVE CRITERIA

Inclusion criteria

- Age 18 or older
- Low cervical or thoracic SCI (C6-T12)
- ASIA Impairment Scale grade A or B for mid to low thoracic injuries (T4-12) or ASIA grades A, B, or C for cervical and high thoracic injuries (C6-T3)
- Time postinjury greater than 6 months
- Intact lower motoneurons
- Absence of psychologic problems or substance abuse
- Range of motion within normative limits
- Full coverage of the acetabulum and minimal laxity of the knee or ankle
- No acute medical problems
- Adequate social support and stability
- Willingness to comply with follow-up procedures

Exclusion criteria

- Pacemaker
- Cardiac arrhythmia
- Pregnancy
- Contractures of upper or lower extremities
- Seizure disorder
- Obesity
- Untreated substance abuse
- Immune deficiency
- Recurrent urinary tract infections
- Presence of pressure ulcer

References


Suppliers

b. Sigmedics Inc, 333 N Broad St, Fairborn, OH 45324.
Expert Opinion and Controversies in Sports and Musculoskeletal Medicine: Concussion in the Young Athlete

Christopher J. Standaert, MD, Stanley A. Herring, MD, Robert C. Cantu, MD, MA


Concussion is a common injury in young athletes and can be very challenging for clinicians to diagnose and manage. Debate exists over not only the incidence of long-term risks of multiple concussions but also the potential for catastrophic outcomes after sports-related head injury. Decisions on returning athletes to competition can be difficult, and there are limited prospective data on which to make these decisions. This has resulted in the existence of a number of published guidelines and consensus statements on the management of concussion in athletes. Athletes sustaining a concussion need appropriate on-field care and structured follow-up. Baseline cognitive assessments can be helpful, but clinicians must be aware that head trauma may result in a wide array of clinical signs and symptoms. Delivery of care and decisions on return to play need to be based on an individual assessment of the affected athlete.

Key Words: Adolescent medicine; Brain concussion; Rehabilitation; Sports medicine.

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

A CONCUSSION IS A MILD traumatic brain injury (TBI) that is induced by a direct or indirect force. The manifestations of concussion can vary widely between patients and may include confusion; amnesia (anterograde or retrograde); disorientation; altered consciousness; headache; dizziness; emotional lability; or a variety of other problems such as those associated with balance, vision, or speech. The pathophysiology of concussion in humans is extrapolated from animal models and involves alterations in glucose metabolism, ionic shifts, altered blood flow, and the release of excitatory neurotransmitters. In young athletes, a potentially lethal condition termed second impact syndrome (SIS), which can occur after a concussion in an athlete still symptomatic from a prior brain injury, is thought to be caused by the loss of autoregulation of cerebral blood flow with subsequent vascular engorgement. Although postconcussion symptoms usually resolve within days to a week, they may persist beyond this and rarely may never entirely clear. Thus, postconcussion symptoms may persist longer than the typical timeframe of game-to-game preparation within a season. Athletes sustaining a concussion, especially in football, are at increased risk for repeat concussion. Substantial debate remains on the long-term consequences of repeated concussions and on how this risk should factor into decisions regarding return to play (RTP).

From a clinical perspective, concussion is a common injury in young athletes and can be extremely challenging for clinicians to manage. Based on their extensive data, Powell and Barber-Foss estimated that over 62,000 high school varsity athletes suffer concussions annually in the United States, representing 5.5% of all reported sports injuries in this population. Injuries associated with football accounted for 63% of the cases, but athletes (both male and female) were affected in a variety of sports. A recent study by Boden et al reviewed 94 incidents of severe head injuries associated with football from the National Center for Catastrophic Sports Injury Research and found that high school players had an incidence rate 3 times that noted in collegiate players. Of the injured athletes for whom telephone interview contacts could be found, 39% had been playing with residual neurologic symptoms from a prior head injury at the time of the catastrophic event. Unfortunately, there are relatively little long-term data on the effects of concussion in young athletes, and management decisions for these athletes can be daunting. The situation is further complicated by the fact that published guidelines are based on adult studies.

CONTROVERSIES

Return to Play

Decisions on both same-day and delayed RTP can be difficult, and identifying the symptoms and neurologic deficits associated with mild TBI is crucial. Debate exists on which findings are most critical in the decision-making process, but all signs and symptoms need to have cleared at rest and with exertion before the athlete is returned to play.

Value of Published Guidelines

There are currently at least 20 published grading scales or guidelines for the management of concussion in athletes. The grading scales vary in the importance assigned to different factors, with a particular point of controversy being the relative importance of loss of consciousness (LOC) versus post-traumatic amnesia (PTA). Only one of the currently available grading scales is evidence based, and none have been evaluated by a double-blinded prospective study of injured athletes.

Risk of Catastrophic Injury or Permanent Cognitive Deficits

The precise incidence of SIS is not known, in part because the number of players participating with postconcussion symptoms is unknown. Similarly, the number of players who have had to terminate careers because of postconcussion syndrome, including permanent cognitive deficits, is not known despite some high-profile professional athletic cases in the media.
There are published data on the cognitive functioning of previously concussed athletes, but the number of studies is small and the results inconsistent. Thus, accurate assessment of risk for these conditions is not possible with current data.

PRACTICAL APPROACHES

How Do I Assess Athletes on the Field?

The first step is to identify a player who may have sustained a concussion. This may be obvious in the case of an overt alteration in consciousness or orientation, but medical personnel need to be alert for more subtle features of confusion, disorientation, or altered physical, cognitive, or emotional status. Initial evaluation of the "downed athlete" requires assessment for airway, breathing, circulation, and level of consciousness. Further assessment of neurologic status and the cervical spine allows for determination of whether the athlete is moved to the sideline or prepared for emergent transport off the field. Once the need for emergent transport has been excluded, the sideline evaluation of the potentially concussed athlete includes a more detailed history and neurologic examination. Cognitive, somatic, and affective findings need to be addressed (appendix A). A structured approach to sideline evaluation is necessary. A number of standardized testing systems, including electronic portable devices, are available for use, such as the Standardized Assessment of Concussion. These systems generally only evaluate cognitive function, and baseline preinjury assessment is required for the clinical assessment of data acquired after a concussion. Particular attention needs to be devoted to PTA, both retrograde and anterograde, because this is the risk factor most predictive of persisting symptoms and cognitive deficits.

Players sustaining mild TBI need to be supervised at all times on the field, with postgame observation clearly defined. Serial neurologic assessments are important. Neurologic deterioration; prolonged LOC (minutes as opposed to seconds); or any decline in level of consciousness, vomiting, severe or worsening headache, seizures, or prolonged amnesia (15–30 min) are among the indicators for emergent transport to a medical facility.

Can the Athlete RTP the Same Day?

Clearly, emergent considerations need to be excluded before same-day RTP is even an option. Athletes with persisting signs and symptoms or any duration of amnesia should not RTP. In other situations, there is substantial disagreement among authors on returning injured athletes to play the same day, with some authors allowing RTP if all symptoms clear rapidly. Even with rapid resolution of symptoms, however, athletes should undergo a physiologic challenge such as running or push-ups to ensure that they remain asymptomatic before resuming competition. Unfortunately, there are no prospective data on this issue. Given the unique metabolic and vascular circumstances in young athletes and what are typically relatively reduced resources for baseline and on-field care, the safest decision, if there is any concern, is to leave the athlete sustaining a concussion out of play. In sum, "when in doubt, sit them out."

How Is the Decision Made on When They Can RTP?

Ideally, a decision to RTP will be based on the evaluation of each athlete. The decision depends on a number of factors, including the extent of injury, the duration of symptoms or cognitive dysfunction, and the history and proximity of previous concussions. There is significant controversy on the means of grading concussion and deciding on RTP. One major published guideline uses LOC as the dominant marker for injury severity, but LOC has since been shown to be poorly predictive of symptoms and cognitive disturbance after a concussion. Athletic concussion is also rarely associated with LOC of more than 1 minute. Although amnesia appears to be more predictive of short-term outcome, the role of ongoing assessment of cognitive and physical functioning of the athlete is unclear. In their study of collegiate football players, McCrea et al noted that the average time to symptom resolution in players sustaining a concussion was 7 days, with cognitive function improving to baseline in 5 to 7 days and balance deficits clearing by 3 to 5 days. The athletes in this study, however, followed varying trajectories with 10% having symptoms for more than a week. In a related study, Guskiewicz et al found that 91% of players studied who sustained a second concussion within the same season did so within 10 days of the initial concussion. These data should give pause to those considering a rapid RTP for concussed athletes, particularly in the absence of baseline testing and structured, serial postinjury neurocognitive and physical evaluation.

Restrictions on RTP generally increase with the severity of concussion, and the most widely referenced guidelines suggest more prolonged time out of competition after a second concussion in the same season and termination of the season or participation in contact sports altogether with a third severe concussion. As noted, additional caution needs to be given to returning the young athlete to play prematurely because of their unique risk profile. A number of authors advise following a progressive, structured RTP protocol. This starts with complete rest (both physical and cognitive) after mild TBI; a gradual progression of activity is then initiated after symptom resolution. Low-level activities are performed initially, and over several days the athlete is progressed through higher-level and sport-specific skills, noncontact drills, and controlled contact before considering RTP. By following the numbers presented, it is clear that if on any given Friday a high school athlete sustains a concussion that does not resolve rapidly, he/she has a very high likelihood of not being cleared to play by the following Friday. A single concussion may thus be expected to result in a multi-game absence. Given the data, it is sobering to note that the median time loss from sports activity after mild TBI in the study of Powell and Barber-Foss was 3 days.

What About Using the Published Guidelines?

Overreliance on published guidelines is problematic for a number of reasons, including the nonuniformity of their guidance, lack of prospective validation in most, the reliance on LOC as the marker of injury severity, and the extent of individual variation of mild TBI. However, it is important to understand the guidelines, and they can provide a reference point from which to assess an injured athlete. The modified Cantu guidelines (by an author of this article) was developed through extensive clinical experience and a thorough assessment of the available evidence and may be the most clinically relevant of the published guidelines. Overall, it is our opinion that the management of athletes sustaining a concussion should be performed on an individual basis by using clinical signs and symptoms and appropriate follow-up as opposed to strict adherence to published grading scales.

What Is the Risk of Recurrent Concussion, SIS, or Permanent Cognitive Deficits?

The data on these issues are variable. Athletes sustaining a single concussion clearly seem to be at an increased risk for a recurrent concussion. SIS is a rare but devastating occurrence that, other than in boxers, has never been reported in anyone older than 19 years of age. The exact risk factors for
this syndrome are unclear, particularly given the low rate of occurrence, but RTP in close proximity to the time of a prior concussion with persisting signs or symptoms is clearly a significant concern when considering the evidence on catastrophic injury as well as cumulative injury. In general, research in this area has been hampered by difficulty with recall bias among subjects, defining "concussion," and obtaining baseline data. Clearly, prospective studies on a large pool of young athletes at risk for mild TBI with long-term follow-up, although difficult, is required to definitively address this question. Until such information is available, clinicians, athletes, parents, and coaches need to be aware of the potential for long-term consequences after a concussion and should proceed on an individual basis.

**How Can I Best Prepare to Manage Athletes at Risk for Concussion?**

There are several ways to manage athletes at risk for concussion including (1) know your athlete; (2) have baseline health and cognitive data whenever possible; (3) establish effective communication with the coaching staff, trainers, athletes, and families; (4) be prepared to manage acute injury; (5) have a system for providing appropriate sideline and postgame supervision of the injured athlete; (6) provide structured, ongoing follow-up; and (7) know the literature.

**APPENDIX 1: SELECTED SIGNS AND SYMPTOMS OF CONCUSSION**

<table>
<thead>
<tr>
<th>Cognitive</th>
<th>Somatic</th>
<th>Affective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confusion</td>
<td>Headache</td>
<td>Emotional lability</td>
</tr>
<tr>
<td>Post-traumatic amnesia</td>
<td>Fatigue</td>
<td>Irritability</td>
</tr>
<tr>
<td>Retrograde amnesia</td>
<td>Disequilibrium</td>
<td></td>
</tr>
<tr>
<td>Loss of consciousness</td>
<td>Dizziness</td>
<td></td>
</tr>
<tr>
<td>Disorientation</td>
<td>Nausea/vomiting</td>
<td></td>
</tr>
<tr>
<td>Feeling “zoned out”</td>
<td>Visual disturbances</td>
<td></td>
</tr>
<tr>
<td>Vacant stare</td>
<td>Photophobia</td>
<td></td>
</tr>
<tr>
<td>Inability to focus</td>
<td>Phonophobia</td>
<td></td>
</tr>
<tr>
<td>Delayed verbal and motor responses</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slurred or incoherent speech</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excessive drowsiness</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Reprinted with permission from the American College of Sports Medicine.1

**References**


Transition to adulthood for youth with developmental disabilities has become an important concern internationally of service providers working with these young people. Reflecting on the useful review by Binks and colleagues in this issue of the Archives, we argue as developmentalists that this is an ideal time to step back from our traditional preoccupation with “treatment” of childhood disability and to reconsider broadly what our goals for intervention ought to be. We invoke the concepts of the International Classification of Functioning, Health and Disability framework and draw on research that taps the voices of young people with disabilities—voices we believe have a lot to tell us about what has and has not worked for them. We suggest that there are unparalleled opportunities to enhance transition to adulthood for young people with developmental disabilities, in part by a feed-forward of the best of childhood services, and to work to prevent many of the difficulties faced by the current generation making this challenging transition.

Key Words: Disabled children; Pediatrics; Rehabilitation.
© 2007 by the American Congress of Rehabilitation Medicine and the American Academy of Physical Medicine and Rehabilitation

In their comprehensive and useful review article on the issues of transition to adulthood of people with developmental disabilities in this current issue of Archives, Binks et al1 provide valuable perspectives on the increasing challenge of making effective service transitions to adult care for youth with special needs. We appreciate the opportunity to comment on this internationally important problem. In doing so, we would like to step back from the review itself and reflect on the broader implications of the transition issues raised by Binks. We will suggest that the whole system of services for children and youth with “developmental disabilities” should be reconﬁgured to anticipate—and ultimately to prevent—many of the problems currently facing youth with special needs, their families, and the “systems” with which they come in contact.

To begin with, we believe that the issues discussed by Binks1 regarding “transition from child-centered to adult-centered health care” are intimately entangled with many of the traditions and assumptions of developmental rehabilitation. In the rehabilitation of adults who develop or acquire functional limitations, the goals of intervention are usually fairly obvious—we are trying to return the person as closely as possible to their previous level of functional capability, a level that was probably well deﬁned for that person before the incapacity occurred. Imported from the world of adult medicine, classical “early interventions” for children with disabilities and functional difﬁculties have had as a core assumption the importance of addressing “impairments” and using “normal” development as the guidepost. Thus, for example, therapy for children with cerebral palsy (CP) has traditionally endeavored to “normalize” increased muscle tone, inhibit reﬂexes that appear to interfere with “normal” function, promote “normal” movement, and so forth. We are ﬁnding, however, that there is little evidence that these apparently useful goals are meaningfully achievable with children with conditions such as CP. Furthermore, studies that explore the assumption that a change in “impairment” will translate into improvements in “activity” or “participation” have led people to question whether this does indeed happen.2

An alternative approach to thinking about and “managing” childhood disability is to recognize formally the notion of disabilities as challenges to children’s typical development (hence, the notion of “developmental disability”). We believe that it is essential to see beyond the biomedical aspects of diagnosis and pathology and to consider, in all that we do with children and families, how we can support and promote child (and family) development, achievement, and the emergence of a sense of self and of capability in the face of challenges to functional achievement presented by the functional differences that are inherent in developmental disability.

There are indications that a more functionally oriented approach to developmental intervention is being embraced, at least in principle, by the pediatric service community. Some child-focused developmental programs provide comprehensive care, continuity, coordination across disciplines, and family engagement (in family-centered service models3)—all elements that are reported by Binks1 to be relatively poorly available in the world of adult services. This suggests that there are lessons to be exported (fed forward) from the child and youth services to what one hopes will become specialty services for adults with a pediatric or lifelong disorder.4

We believe that the field of childhood disability is on the brink of a major paradigm shift, of which the transition challenges represent an important component. With the advent and rapid adoption of the concepts of the World Health Organization’s International Classiﬁcation of Functioning, Health and Disability (ICF),5 the scope of developmental services is beginning to expand its horizons.6 Moving beyond concerns with “body structures and functions,” professionals are promoting function through an emphasis on activity, encouraging participation in society, and acknowledging the role of personal factors and environmental factors as legitimate components of intervention programs. At the same time, the considerable interest in quality of life (however deﬁned) is prompting people to expand their focus beyond functional disability in an effort...
to address issues important to the people who have the conditions for which services are provided.

Readers may be wondering at this point how these musings about childhood disability relate to transition to adulthood. As we see it, the identification and discussion of the challenges facing youth and young adults with developmental disabilities, as identified by both Binks and ourselves in this commentary, should force child-oriented service providers to reflect on what youth need as they make the transition from pediatric to adult services. This, in turn, begs the question of what we are trying to achieve with our interventions during the childhood years to help prepare young people and their families for transition. What are our assumptions about disability? What do we think about the needs of children and families? How do we set intervention goals, and whose goals should they be? Where does family engagement and responsibility fit into this mix of issues? Also, what component of our orientation in interventions is toward “fixing” and how much is directed toward healthy functional development?

Answers to these questions are coming directly from youth with developmental disabilities themselves as they grow up, and we believe that all service providers should be listening to them. Feedback from adults who have made the transition is consistent in telling us that our orientation to impairment management is not what they believe they needed in their developing years. Evidence supports the importance of addressing their functional needs by using an individualized approach across the developmental continuum. Furthermore, a focus on functional needs of children within their daily environments, as identified by their families (and by the older children), appears to produce better outcomes, with less therapy.

If we are to refocus our intervention toward personal developmental of children with disabilities (and the development of their families), how might this be done and what would such interventions look like? We believe that in addition to the acquisition of functional skills across the broad range of self-care, mobility, and cognitive abilities, it is essential that we promote opportunities for young people to develop a sense of self including self-confidence, self-esteem, self-determination, and self-management, along with problem-solving and decision-making skills and experiences. They need to acquire a good understanding of their life circumstances and learn how to be articulate about their life needs (beyond simply their “impairment” needs). They need opportunities to participate in real-life experiences throughout childhood and adolescence such as playing in their neighborhood, doing household chores, and having a part-time job. Their families need support, from the outset, to be in control of their agendas and see the disability component of their children’s lives in the broader context of development toward directing their own lives. And, both youth and families need relevant information to enable them to make informed decisions about the future. Evidence-based information-management tools like the CanChild K.I.T. (Keeping It Together) are available to promote a sense of parent competency, with similar tools and resources (eg, the Youth K.I.T. and the DOOR). 2 Adulthood website in Ontario, Canada) under development for use by youth themselves.

Recent evidence also supports the need to focus efforts on changing the environment, to reduce barriers and build supports as much as, and perhaps more than, focusing on skill development of the young person. This includes looking at what health care services and providers are doing. Binks reports that the literature identifies “the inability of child-centered health care professionals (eg, pediatricians) to ‘let go’ of their long-standing relationships with patients” as an important barrier to effective transitions. We wonder whether this is too strong an indictment of both child-oriented practitioners and their youthful patients and families, who may equally not want to sever what have become comfortable and effective relationships because of what are perceived to be inadequately responsive “adult” services. Most adult services appear to be fragmented and unprepared for young adults with developmental disabilities who are living longer lives and expecting to function as full citizens in society. A broad-based environmental approach is needed to make changes in adult services to ensure that these young people experience a successful transition. This also speaks to the need to develop environmental strategies within the adult world as a whole, such as community capacity building, to promote inclusion and participation of all young adults.

There has probably never been a better time in the field of childhood disability to reconsider and improve the way we do our work. If we are indeed in the midst of a paradigm shift and there is “instability” in the system, this is exactly the time to challenge all our traditional assumptions about childhood disability and to embrace exciting new opportunities to consider a broader array of activities directed toward child, youth, and family development. Rather than throwing out what is best in our traditions of developmental intervention, we must take the opportunity to embrace and add concepts inherent in the ICF framework, such as participation and environmental factors, and to be concerned about people’s quality of life, by thinking developmentally and not just in biomedical terms. If we do so, we can promote the best development and participation of young people who grow up with disabilities and hope that in a decade follow-up review articles like those of Binks will report a much improved outlook for the service transitions experienced by young people with developmental disabilities.

We thank Binks for synthesizing the issues so effectively and challenging all of us to reflect on how we can do a more effective job of “transition to adulthood” for youth with developmental disabilities.

References

2. Wright V. How do changes in impairment, activity and participation relate to each other? Results of a study of a group of young people with cerebral palsy who have received lower extremity botulinum toxin type-A injections [PhD dissertation]. Hamilton: McMaster Univ; 2005.
4. Jahnsen R. Being adult with a “childhood disease”—a survey on adults with cerebral palsy in Norway [PhD thesis]. Oslo: Section for Health Science, Faculty of Medicine, Univ Oslo; 2004.


Letters to the Editor

Consortium for Spinal Cord Medicine

We were stunned to discover that an article in the March 2007 supplement includes a key reference identified as a publication by the Consortium for Spinal Cord Medicine (CSCM) entitled Early Acute Management in Adults With Spinal Cord Injury: A Clinical Practice Guideline for Health-Care Providers published in 2007 by Paralyzed Veterans of America. Simply stated, this reference does not exist. A document by this title is presently in draft format only. This draft document is confidential and has not been through field review, legal review, or approved for publication by the CSCM’s Steering Committee.

Any statements attributed to this unpublished document must be recognized as unsupported, and do not represent the opinion or position of the CSCM. The CSCM’s guideline development process occurs in deliberate stages of review ensuring scientific and legal integrity of the field, CSCM participants, the development panel, and Paralyzed Veterans. This same process has resulted in successful publication of 9 previous titles with 6 companion consumer guides. Any breach of this process subjects the entire process to suspicion and criticism, which would undermine the collaborative enterprise of this Consortium of 22-member organizations.

Lawrence Vogel, MD
CSCM Steering Committee Chair

Early Acute Topic Guideline Development Panel Chair

Peter Wing, MD
Paralyzed Veterans of America

Washington, DC

Reference

doi:10.1016/j.apmr.2007.06.766

The author responds

During the preparation of the Spinal Cord Injury Medicine section of the 2007 Study Guide (published as a supplement to the March issue of Archives of Physical Medicine and Rehabilitation), I was participating on a panel for the Paralyzed Veterans of America’s Consortium for Spinal Cord Medicine (CSCM) in the development of a document entitled Early Acute Management in Adults With Spinal Cord Injury: A Clinical Practice Guideline for Health-Care Providers. Anticipating publication of the CSCM guideline around the same time as the Study Guide, I included in the Study Guide (principally in the article on which I was the lead author) references that anticipated the CSCM document. In error, I allowed my work on the Study Guide to proceed to publication without removing (or correcting) references to the incomplete and unpublished CSCM document. The culpability for this error rests solely with me, as neither the CSCM, the Paralyzed Veterans of America, nor the development panel was aware that the Study Guide included such a reference. Indeed, the editor of this edition of the Archives was also unaware that the CSCM document was not pending publication.

I apologize for my error and failure to correct the Study Guide prior to publication and direct readers to the accompanying correction.

Lisa-Ann Wuermser, MD
Mayo Clinic
Rochester, MN

Reference

doi:10.1016/j.apmr.2007.06.767

Corrections


Reference 2 (Consortium for Spinal Cord Medicine. Early acute management in adults with spinal cord injury: a clinical practice guideline for health-care providers. Washington (DC): Paralyzed Veterans of America; 2007) should be deleted (currently a draft document is in field review, with publication expected in late 2007), and all text referencing this document should be deleted or reference an existing publication. Listed below is each instance this unpublished document is referenced in the text, with a correction, an alternative reference, or both.

In section 2.1, paragraph 1 (pS55), the sentence “Transfer to such a center is advocated as soon as the patient is stable, with the suggestion that emergency medical services in urban areas should consider bypassing the nearest hospital to take SCI patients to level 1 trauma centers directly” should be deleted; the sentence cannot be supported by existing literature alone.
In section 2.1, paragraph 3 (pS55), the sentence that begins “Both the neurosurgical guidelines" and the Consortium for Spinal Cord Medicine clinical practice guidelines consider the use of high-dose methylprednisolone to be a treatment option rather than a standard” should be changed to read: “The American Academy of Neurological Surgeons and Congress of Neurological Surgeons clinical practice guideline for acute cervical SCI considers the use of high-dose methylprednisolone to be a treatment option rather than a standard. Reference 2 should be deleted and reference 12 should be corrected to: Pharmacologic therapy after acute cervical spinal cord injury. Neurosurgery 2002;50(3 Suppl):S63-72.

In section 2.1, paragraph 5 (pS56), the sentence “Sepsis, hypovolemia, and cardiogenic shock must all be considered in the early period of SCI” should be deleted. Although the statement makes common sense, the literature does not support this specific statement.

In section 2.1, paragraph 6 (pS56), the source for the statement “Preventing complications remains vital in acute SCI. Prophylaxis of venous thromboembolism should begin no later than 72 hours after onset of SCI and should include anticoagulation for most injuries” should be references 22 and 23. Reference 2 should be deleted.

In section 2.1, paragraph 7 (pS56), the source for the statement “Either H2-blockers or proton pump inhibitors are indicated, to be started at admission and continued for 4 weeks” is Lu WY, Rhoney DH, Boling WB, Johnson JD, Smith TC. A review of stress ulcer prophylaxis in the neurosurgical intensive care unit. Neurosurgery 1997;41:416-26. Reference 2 should be deleted.

In section 2.2, paragraph 3 (pS56), sentence 2 should be revised to read: “A bowel program should be initiated in acute care and can reasonably begin soon after enteral feeding is initiated.” Reference 29 is: Consortium for Spinal Cord Medicine. Neurogenic bowel management in adults with spinal cord injury. Washington (DC): Paralyzed Veterans of America; 1998. Reference 2 should be deleted.

In section 2.2, paragraph 3 (pS56-7), sentence 2 should be deleted from the end of the sentence “The removal of an indwelling urinary catheter and initiation of intermittent catheterization can be recommended as soon as the patient no longer requires intravenous fluids and the medical status does not require strict monitoring of urinary outputs.” The following new reference should be added to the next sentence (“However, physiatrists should be aware of the diuresis of third spaced fluid associated with mobilizing the patient and warn against volumes higher than 500mL per catheterization”): Consortium for Spinal Cord Medicine. Bladder management for adults with spinal cord injury: a clinical practice guideline for health-care professionals. Washington (DC): Paralyzed Veterans of America; 2006.

In section 2.2, paragraph 6 (pS57), the last 2 sentences should be corrected to read: “However, low volume ventilation is well established to improve outcomes in the general trauma setting, primarily because of the high frequency of acute lung injury (ALI) and adult respiratory distress syndrome (ARDS) in this population. Therefore, ALI and ARDS should likely be resolved or ruled out prior to initiation of a high volume weaning protocol.” Delete reference 2 at the end of the sentence and add the following new reference (where indicated): Petrucci N, Iacovelli W.Ventilation with lower tidal volumes versus traditional tidal volumes in adults for acute lung injury and acute respiratory distress syndrome. Cochrane Database Syst Rev 2003;(3):CD003844.

In section 2.2, paragraph 7 (pS57), the entire paragraph should be edited to read:

Dysphagia is also a contributor to respiratory deterioration after SCI. Although up to 30% of patients with tetraplegia are found to have dysphagia at admission to inpatient rehabilitation, up to two thirds of patients undergoing elective cervical spine surgery demonstrate postoperative dysphagia, and has been demonstrated in both anterior and posterior cervical spine approaches. It is likely, then, that the rate of dysphagia in acute SCI is higher than that documented later in their course in the rehabilitation setting. Persons with cervical spine surgery, tracheostomy, halo stabilization, and older age have been shown to have an increased risk of dysphagia. In the presence of a rigid cervical orthosis or halo, it is difficult to impossible to compensate for dysphagia using a typical chin-tuck position. Concomitant brain injury is also a risk factor for dysphagia. Use of nasogastric access for enteral feeding can be used until the risk of aspiration diminishes.


In section 2.2, paragraph 9 (pS57), the final sentence “Once the spine is stable, routine turning every 2 hours should be implemented” should reference existing note 43 (Consortium for Spinal Cord Medicine. Pressure ulcer prevention and treatment following spinal cord injury: a clinical practice guideline for health-care professionals. Washington (DC): Paralyzed Veterans of America; 2000.).

In addition, the question and answer sections include 2 references to the unpublished work.
In the 2007 SAE-P: Spinal Cord Injury Medicine section, question 6 (pS89), the text of the question needs no revision. The reference to the unpublished Consortium for Spinal Cord Medicine (CSCM) document (reference a) should be deleted. In its place should be the following 2 references, which were referenced in the CSCM document:


In the answer key to question 6 (pS93), the text should cite the American Academy of Neurological Surgeons and Congress of Neurological Surgeons clinical practice guideline (referenced above) not the CSCM document.

In the 2007 SAE-P: Spinal Cord Injury Medicine section, question 8 (pS89-90), the text of the question needs no revision. The 3 incorrect answers are supported by existing literature (as noted above). Both references (reference a and b) should be deleted. In their place should be the following: Ref: Clinical Activity 2.2.
**ARTICLE OBJECTIVES:**

**Article One:** Circuit Class Therapy Versus Individual Physiotherapy Sessions During Inpatient Stroke Rehabilitation: A Controlled Trial

Learning Objective: After completing this article and with appropriate self-study, the participant will be able to:

a) Describe the components of the circuit training and compare the components to the specifics of individualized physical therapy.

b) Describe the difference in the primary outcomes between the 2 groups studied and between discharge and follow-up.

c) Analyze the differences in the circuit training group and the individual therapy group as they relate to walking independently.

**Article Two:** Role of Sonographic Examination in Traumatic Knee Internal Derangement

Learning Objective: After completing this article and with appropriate self-study, the participant will be able to:

a) Describe the appearance of a knee effusion on ultrasound.

b) List 3 factors that may influence the correlation between knee effusion and internal derangement.

c) Summarize the patients with false-negative and false-positive findings on ultrasound.

**Article Three:** Learning Effects of Repetitive Administrations of the Sensory Organization Test in Healthy Young Adults

Learning Objective: After completing this article and with appropriate self-study, the participant will be able to:

a) Recall the 3 main sensory systems involved in balance.

b) List the 6 conditions of the Sensory Organization Test (SOT).

c) Discuss the improvement in the composite score of the SOT and of the individual components over repeated testing and the relevance of this improvement.

**INDICATE THE DEGREE TO WHICH YOU AGREE OR DISAGREE WITH EACH STATEMENT**

<table>
<thead>
<tr>
<th></th>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Not Certain</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Objectives for article one were met.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Objectives for article two were met.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Objectives for article three were met.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. I learned a new skill or patient management approach.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. This material will enhance my professional effectiveness.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. I plan to implement a change(s) to my practice as a result of this material.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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>AAPM &amp; R/ACRM member</th>
<th>$20.00</th>
</tr>
</thead>
<tbody>
<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

Mitch worked tirelessly to help patients and families dealing with the consequences of TBI. To this end, he served on the boards of 5 state brain injury associations, and was a founding board member of the Brain Injury Association of America (BIAA).

He was repeatedly recognized by his peers as a leader and major contributor to the field. He was elected as a fellow of the American Psychological Association (APA), in both clinical neuropsychology and rehabilitation psychology. He was particularly proud of being selected by the BIAA to receive the first Sheldon Berrol, MD, Clinical Service Award, established to memorialize the important contributions of his friend and colleague, Shelly Berrol. The University of Texas Medical Branch at Galveston chose him to be the first recipient of the Robert L. Moody Award for Distinguished Initiatives in Brain Injury Research and Rehabilitation in 2001.

The American Congress of Rehabilitation Medicine was a major beneficiary of Dr. Rosenthal’s energies. He served on 12 Congress committees, and then, as a board member, and as treasurer, vice-president, and president. The Congress recognized his efforts by awarding him the prestigious Gold Key Award in 2002.

The profession of psychology also benefited greatly from his work. He served as a trustee of both the American Board of Professional Psychology and the American Board of Rehabilitation Psychology. In addition, he held many committee positions and later served as president of the Division of Rehabilitation Psychology of the APA in 1992.

Mitch’s many other accomplishments and contributions are literally too numerous to mention at this time. But all of them do not tell his whole story as a colleague, husband (to Peggy), father (to Michelle and David), and friend. Mitch was dedicated to our field and his profession, but he was also a very special person. He was a survivor of cancer for 27 years, a fact probably little known by many. He was internationally recognized as a lecturer and speaker in TBI. He was highly regarded for his insight and clear communication abilities. And personally, he was a joy to be with. His sense of humor was widely appreciated. He had a sense of irony and wit that did not jest at the expense of others, and enjoyed being occasionally compared with the character of George Costanza from the television sitcom Seinfeld.

Mitch Rosenthal was a Good Man. He had Good Intentions, Good Judgment, and Good Humor. With his passing, there is less good in our field, and we all have a responsibility to try at least a bit harder, to fill the void that his passing has created. We will all miss him dearly, but cherish our memories of him.

Bruce M. Gans, MD
Kessler Institute for Rehabilitation
West Orange, NJ

Arch Phys Med Rehabil Vol 88, August 2007