
Research Article
Medicare Issues

169 The Effect of Medicare's Prospective Payment System on Patient Satisfaction: An Illustration with Four Rehabilitation Hospitals.
Parag K. Shah, MD; Allen W. Heinemann, PhD, ABPP; Larry M. Manheim, PhD

176 ADL Assessment of Nondisabled Turkish Children with the WeeFIM Instrument.
Canan Aybay, MD; Gulten Erkin, MD; Atilla H. Elhan, PhD; Hulya Sirzai, MD; Sumru Ozel, MD

183 The Myofascial Trigger Point Region: Correlation Between the Degree of Irritability and the Prevalence of Endplate Noise.
Ta-Shen Kuan, MD, MS; Yueh-Ling Hsieh, PhD, RPT; Shu-Min Chen, MD; Jo-Tong Chen, MD, PhD; Wei-Chang Yen, MD; Chang-Zern Hong, MD

190 Evaluation of Two Assessment Tools in Predicting Driving Ability of Senior Drivers.
Michael F. Oswanski, MD, FACS; Om P. Sharma, MD, FACS; Shekhar S. Raj, MBBS; Leslie A. Vassar, OTR/L; Kathryn L. Woods, OTR/L; Wendi M. Sargent, OTR/L; Robyn J. Pitock, OTR/L

Book Review

199 Total Knee Arthroplasty.
John A. Schuchmann, MD

Research Article
Spinal Cord Injury

200 A Washing Toilet Seat with a CCD Camera Monitor to Stimulate Bowel Movement in Patients with Spinal Cord Injury.
Ken Uchikawa, MD; Hidetoshi Takahashi, MD, PhD; Genbu Deguchi, OTR; Meigen Liu, MD, PhD

Book Review

204 Anatomical Guide for the Electromyographer: The Limbs and Trunk.
Ralph M. Buschbacher, MD

Research Article
Amputee

Pamela Gallagher, PhD; Olga Horgan, PhD; Franco Franchignoni, MD; Andrea Giordano; Malcolm MacLachlan, PhD

216 Noninvasive Ventilation in Children with Spinal Muscular Atrophy Types 1 and 2.
Albino Petrone, MD; Martino Pavone, MD; Maria B. Chiarini Testa, MD; Francesca Petreschi, MD; Enrico Bertini, MD; Renato Cutrera, MD, PhD
Commentary
Pulmonary

222 Avoiding Respiratory Failure in Neuromuscular Disease: Why Is It Not Done?
   John R. Bach, MD

Invited Review
Nanotechnology

225 Introduction to Nanotechnology: Potential Applications in Physical Medicine and Rehabilitation.
   Assaf T. Gordon, MD; Greg E. Lutz, MD; Michael L. Boninger, MD; Rory A. Cooper, PhD

Case Report
Spinal Cord Injury

242 Spinal Cord Injury Associated with Thoracic Osteoporotic Fracture.
   Sibel Özbudak Demir, MD; Ceyda Akn, MD; Meltem Aras, MD; Füsun Köseolu, MD

Letters to the Editor

247 RE: RACIAL DISPARITIES IN ACCESS TO CARDIAC REHABILITATION.
   Nitin B. Jain, MD, MSPH; Jeffrey N. Katz, MD, MS

247 RE: RACIAL DISPARITIES IN ACCESS TO CARDIAC REHABILITATION.
   Patricia C. Gregory, MD; Thomas A. LaVeist, PhD; Crystal F. Simpson, MD, MHS

Visual Vignette

250 A Technique for Ultrasound-Guided Intrathecal Drug-Delivery System Refills.
   Mark-Friedrich B. Hurdle, MD; Adam J. Locketz, MD; Jay Smith, MD
The Effect of Medicare’s Prospective Payment System on Patient Satisfaction
An Illustration with Four Rehabilitation Hospitals

ABSTRACT


Objective: To examine the impact of Medicare’s Prospective Payment System (PPS) on patient satisfaction at four inpatient rehabilitation hospitals.

Design: Prospective study using a satisfaction survey to examine the effects of Medicare’s PPS for rehabilitation hospitals. Surveys were conducted at four affiliated rehabilitation hospitals in the Midwest.

Results: Patient characteristics varied only slightly pre- to post-PPS, and several characteristics were related to overall satisfaction, including motor functional gain, discharge to home, and respondent (patient or proxy). A 12-point increase on a 12-item motor function scale resulted in 1.13 greater odds (95% CI: 1.04, 1.24) of reporting excellent satisfaction. Patient respondents were 1.27 times more likely (95% CI: 1.07, 1.50) than proxies to report excellent satisfaction, and patients discharged home were 1.65 times more likely (95% CI: 1.31, 2.07) to report excellent satisfaction than patients discharged elsewhere. We found an increase in observed satisfaction from 60.3 to 63.4% (P < 0.05) after PPS implementation, despite a decrease in motor FIM gain.

Conclusions: Patient characteristics such as motor FIM gain, discharge status, and respondent type were significantly associated, although only slightly, with patient satisfaction in inpatient rehabilitation. Percentage of excellent satisfaction improved at these four facilities after PPS implementation, despite declines in motor FIM gain. The improvement may be the result of numerous ongoing quality-improvement initiatives directed at improving patient satisfaction at these facilities.

Key Words: Patient Satisfaction, Rehabilitation, Prospective Payment System, Health Policy
In recent years, many changes have occurred in the nature of reimbursement and payment for inpatient rehabilitation facilities, commonly known as rehabilitation hospitals. The most significant change was the implementation of a Prospective Payment System (PPS) by the Centers of Medicaid and Medicare Services in 2002. The new payment plan pays facilities a specific amount per admission on the basis of a patient’s diagnosis, age, functional status, and comorbidities. This type of reimbursement plan creates incentives to reduce length of stay and amount of services to minimize costs to the facility. Prior studies about the effect of PPS for patients with spinal cord injury have shown that average length of stay decreased in response to PPS.1 Given potential decreases in length of stay and services, one might expect that other outcome measures would also be affected. However, many facilities continuously implement programs as part of their quality-improvement initiatives. These programs are often begun in response to a drop in a certain area of satisfaction or in response to an anticipated change in care, such as a decreased length of stay. These programs may influence the effect of changes in length of stay or services rendered brought on by PPS. For this reason, the changes brought on by PPS may not affect satisfaction in the anticipated manner. In this paper, we examine the effect of PPS on patient satisfaction in four diverse rehabilitation hospitals.

Patient satisfaction is well established as an important outcome in health care.2–6 From a clinical perspective, satisfied patients cooperate more with treatment regimens and achieve better clinical outcomes.7 From a business perspective, satisfied patients are more likely to be loyal to treatment providers and less likely to bring malpractice suits.8 Patient satisfaction is also identified as a measure of healthcare quality,9,10 and recent efforts have focused on making these data available to consumers. A 2001 Institute of Medicine report recommended publicly reporting patient-satisfaction data as a measure of quality.11 The Rhode Island state legislature mandated reporting patient-satisfaction data to aid consumers in selecting healthcare services,12 and two government agencies are developing a survey for public reporting of patient satisfaction to aid consumer choice and to measure healthcare quality.13 Aside from aiding consumers, measuring patient satisfaction is also important for facilities and for the industry as a whole. The ability to compare data for varying populations would enable hospitals to evaluate accurately the effect of a quality-improvement initiative or other intervention and to compare one hospital with other facilities.

To effectively evaluate the effect of an intervention on patient satisfaction, one must control factors, such as patient characteristics, that may influence patient satisfaction. To normalize populations and detect changes over time, studies of patient satisfaction in various types of healthcare settings have employed control charts, risk adjustment, and raw-score analysis.14–16 In this study, we compare raw scores with risk-adjusted, or predicted, scores. Risk adjustment is necessary because different kinds of patients may respond differently to satisfaction questions independently of the care they receive.17 To isolate the effects of a policy change that may concurrently affect quality and change patient sociodemographics, patient characteristics must be considered. Previous work on the effect of risk adjusting for different patient and clinical factors has been equivocal. Several studies have shown satisfaction is related to age and clinical status, whereas other studies have not.4,6,18 Prior studies have demonstrated that patient characteristics account for 2–7% of the variation in satisfaction ratings.17,19,20

The patient and clinical characteristics that affect satisfaction in other venues of healthcare delivery may not be the same in rehabilitation care.21 Rehabilitation patients are recovering from life-changing health events and learning how to rebuild their lives with skills and abilities that have changed markedly. Clinician expectations are often not for full functional recovery but for rebuilding a satisfying and meaningful life. To learn self-sufficiency, management of chronic diseases, and management of daily activities with a disability, satisfaction with the care a patient is receiving is of critical importance.3,18 A few studies have examined patient characteristics and their relationship to patient satisfaction in rehabilitation facilities. Franchignoni et al.1 have reported that functional status was unrelated to satisfaction of rehabilitation inpatients. This study was conducted at one Italian rehabilitation center and included only 55 patients. In another series of articles that explored the relationships between demographic and functional status and satisfaction, functional gains and discharge functional status were found to have a significant effect on satisfaction in specific rehabilitation populations.22–25 Specifically, for stroke patients, proxy and patient respondents reported satisfaction differently, with proxies reporting lower satisfaction scores. Furthermore, persons showing cognitive and motor gain were more likely to be dissatisfied than patients showing no gain.23 The discharge setting was a significant predictor in a study of patients with orthopedic injury, with those patients discharged home about half as likely to be dissatisfied.24 Another study observed that increased FIM discharge scores were significant
predictors of higher satisfaction in patients with any type of cerebrovascular injury.25

In light of this limited and inconsistent information, we examined three questions: (1) To what extent are patient characteristics related to satisfaction? (2) To what extent does adjusting for patient characteristics change satisfaction scores after an intervention that might change satisfaction along with patient characteristics? (3) What are the effects of PPS on patient satisfaction in a sample of diverse rehabilitation hospitals? We report raw and predicted scores from patients 1 yr before and 1 yr after implementation of PPS; then, we apply risk-adjustment techniques, and then we examine changes in satisfaction.

METHODS

Recruitment

The four sites from which patient data were analyzed are all affiliated with a midwestern network of inpatient and outpatient rehabilitation centers and have distinct case-mix characteristics. The four rehabilitation hospitals included a free-standing urban hospital with 155 beds (rehabilitation hospital A), a 35-bed suburban unit (rehabilitation hospital B), a 32-bed unit in a rural community (rehabilitation hospital C), and a 24-bed unit within a level I trauma center (rehabilitation hospital D). By selection, all patients were over the age of 18, had a length of stay greater than 4 days, and were alive 1 mo after discharge. For the pre- and post-PPS comparison, 3276 patients discharged between January 2000 and December 2001 were compared with 4806 patients discharged between January 2003 and December 2004. Data from 2002 were omitted because that was the year that PPS was implemented and a time of transition for rehabilitation hospitals. The questionnaire was administered by telephone to patients 1 mo after discharge. The overall response rate was 55%. Basic demographic characteristics including age, gender, and admission FIM scores were similar between the respondents and nonrespondents. There was no selection or exclusion of any particular medical condition. The subjects closely approximated the overall prevalence of conditions in the chosen rehabilitation hospitals, with approximately 25% with lower-extremity joint replacements and 16% with stroke.

Dependent Variable

The patient-satisfaction instrument used in the health system contains 18 items that focus on various aspects of patient experience and was developed and validated for rehabilitation inpatients. It demonstrates high internal consistency and face validity.6 The questionnaire is part of an ongoing quality-assurance initiative administered by the rehabilitation hospital. For this reason, IRB approval was necessary only for the secondary analysis, not for data collection. All items are rated on a four-point rating scale that ranged from “excellent,” to “good,” “fair,” and “poor.” For analysis, we selected a single item, “overall satisfaction with care,” because it correlates strongly with the total score and reflects patients’ and proxies’ global perceptions. To compute each facility’s satisfaction score, we used the percentage of patients reporting excellent satisfaction as our outcome. Percent “excellent” is a commonly reported measure and a goal among outcomes managers who strive for excellent ratings within their institutions. Also, because satisfaction surveys generally show high levels of satisfaction, differentiating “excellent” from all other answers may reduce some of this bias.

Independent Variables

Risk factors that might affect satisfaction were selected on the basis of a review of the literature and clinical experience. These included respondent type (patient vs. proxy), age, gender, functional gain, and discharge destination. Functional status was measured with the FIM,26 a widely used instrument that is incorporated in Medicare’s patient-assessment inventory for rehabilitation hospitals. We used motor and cognitive subscores of the FIM because they represent two distinct aspects of function.27 The motor FIM subscore is the sum of 12 item scores, each rated on a scale of 1–7. For the motor FIM gain score, we omitted tub transfer because the patient-assessment inventory for inpatient rehabilitation facilities does not use this item.28 The cognitive FIM subscore is the sum of five items, also on a 1–7 scale. The FIM scores were divided by the number of items in each category to construct a modified motor (motor FIM/12) and cognitive (cognitive FIM/5) FIM score. By constructing modified scores, the coefficients generated by the regression models indicate the effect of a 12-point change in the motor score and a 5-point change in the cognitive score, or an average of a 1-point change in each item within the subscore.

According to Carter and Paddock,29 reported FIM scores under PPS are likely to be lower because of changes in scoring rules, as opposed to true changes in patients’ functional status. Scoring changes under PPS altered the way the admission FIM was obtained. Previously, the FIM score measured at one point during the first 72-hr period was used as the admission FIM. Post-PPS scoring changes included measuring the FIM score at multiple times and taking the lowest FIM score within the patient’s first 72 hrs of admission. By conducting multiple measurements, average admission scores would be expected to go down without there being an actual change in functional status. Fur-
thermore, scoring rules changed for certain items such as eating and bowel and bladder function, causing these scores to also drop artificially. To remove pre-PPS to PPS coding differences, we used a pre-PPS to PPS FIM crosswalk developed using Rasch analysis. The pre-PPS FIM scores were adjusted to be equivalent to post-PPS scores according to the crosswalk and were then used in the model.

Statistical Analysis
Risk adjustment of the effect of PPS on patient satisfaction was achieved using a logistic regression analysis with forced entry of predictors to determine odds ratios and to evaluate the significance of risk factors. We generated a model using all pre-PPS patients; then, using the model, we computed a predicted probability of reporting excellent satisfaction for each patient after PPS. We compared the pre-PPS scores with the post-PPS observed scores to measure actual change in satisfaction. Second, we compared the post-PPS predicted scores with post-PPS observed scores to evaluate the magnitude of risk-adjustment effects.

RESULTS
Table 1 shows how patient demographic characteristics and functional gains varied across sites and from before to after PPS. Individual sites varied somewhat in their patients’ characteristics. The rural rehabilitation hospital (rehabilitation hospital C) had a lower percentage of male respondents (32%), and the mean age of patients at the free-standing rehabilitation hospital (rehabilitation hospital A) was lower than the others. The percent discharged home changed significantly at only one site and did not change significantly overall. Mean age increased slightly (66.5 vs. 67.7, \( t = -3.12, df = 8074, P < 0.01 \)), as did the percentage of proxy respondents after PPS (29.4 vs. 31.5%, \( \chi^2 = 4.1, P < 0.05 \)). After adjustment, motor FIM gain decreased after PPS in all but one of the sites. Contrary to expectations, cognitive FIM gain increased in all facilities after PPS.

Table 2 shows odds ratios for the facilities and patient characteristics for the logistic regression model. Factors that significantly affected satisfaction scores included motor FIM gain, respondent, and percentage discharged to home. Patient respondents and patients discharged to home had higher odds ratios (1.27 and 1.65, respectively) compared with proxy respondents and patients discharged to other facilities. A 12-point motor FIM gain was associated with a 1.13-fold greater odds of reporting excellent satisfaction. The suburban and rural facilities had higher odds of reporting excellent satisfaction than the urban facility (3.50 and 1.65, respectively).

Observed and predicted scores are reported in Table 3. Overall observed satisfaction scores increased slightly from 60.3 to 63.4% (\( t = -2.78, df = 8080, P < 0.01 \)). Predicted scores were calculated using the logistic regression model reported in Table 2. The variance explained by the model was low (\( R^2 = 4\%) \); therefore, the predicted score did not differ much from the observed score (61.3 vs. 63.4%). The slight decrease was attributable to the decrease in patient respondents and the decrease in motor FIM gain after PPS. Most facilities, with the exception of the rural facility, had higher post-than pre-PPS scores.

### Table 1: Pre– and post–Prospective Payment System patient characteristics

<table>
<thead>
<tr>
<th>Site</th>
<th>Pre Respondent</th>
<th>Post Respondent</th>
<th>Pre Motor FIM Gain</th>
<th>Post Motor FIM Gain</th>
<th>Pre Cognitive FIM Gain</th>
<th>Post Cognitive FIM Gain</th>
<th>% Discharge Home</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rehabilitation hospital A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>1539</td>
<td>44.5</td>
<td>62.8 (0.48)</td>
<td>71.0</td>
<td>17.5 (0.29)</td>
<td>1.2 (0.90)</td>
<td>88.5</td>
</tr>
<tr>
<td>Post</td>
<td>2239</td>
<td>43.6</td>
<td>63.2 (0.37)</td>
<td>65.3*</td>
<td>18.9 (0.26)*</td>
<td>3.4 (0.11)*</td>
<td>87.4</td>
</tr>
<tr>
<td>Rehabilitation hospital B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>965</td>
<td>40.3</td>
<td>69.5 (0.43)</td>
<td>66.0</td>
<td>20.1 (0.26)</td>
<td>1.6 (0.10)</td>
<td>86.2</td>
</tr>
<tr>
<td>Post</td>
<td>1355</td>
<td>41.7</td>
<td>72.2 (0.33)*</td>
<td>70.0*</td>
<td>14.2 (0.24)*</td>
<td>1.8 (0.10)</td>
<td>83.8</td>
</tr>
<tr>
<td>Rehabilitation hospital C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>375</td>
<td>32.3</td>
<td>71.3 (0.74)</td>
<td>79.5</td>
<td>23.1 (0.51)</td>
<td>1.3 (0.16)</td>
<td>91.2</td>
</tr>
<tr>
<td>Post</td>
<td>756</td>
<td>32.5</td>
<td>71.9 (0.47)</td>
<td>71.3*</td>
<td>22.6 (0.32)</td>
<td>3.7 (0.19)*</td>
<td>94.8*</td>
</tr>
<tr>
<td>Rehabilitation hospital D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>397</td>
<td>37.5</td>
<td>69.5 (0.70)</td>
<td>72.0</td>
<td>22.6 (0.43)</td>
<td>1.5 (0.17)</td>
<td>89.4</td>
</tr>
<tr>
<td>Post</td>
<td>456</td>
<td>45.0*</td>
<td>69.1 (0.69)</td>
<td>75.4</td>
<td>17.3 (0.43)*</td>
<td>2.5 (0.23)*</td>
<td>86.4</td>
</tr>
<tr>
<td>Total</td>
<td>3276</td>
<td>41.0</td>
<td>66.5 (0.29)</td>
<td>70.6</td>
<td>19.5 (0.18)</td>
<td>1.4 (0.06)</td>
<td>88.3</td>
</tr>
<tr>
<td>Post</td>
<td>4806</td>
<td>41.5</td>
<td>67.7 (0.23)*</td>
<td>68.5*</td>
<td>17.9 (0.16)*</td>
<td>2.9 (0.07)*</td>
<td>87.4</td>
</tr>
</tbody>
</table>

*Boldface, pre–post comparison with \( P < 0.05 \).
DISCUSSION

Rehabilitation lengths of stay have decreased in recent years, and it is expected that this trend will continue with the implementation of PPS. An expected consequence of shorter stays is that patients’ functional gains will decline along with patient satisfaction. However, in this study we found results to the contrary. Both observed and predicted satisfaction increased after the implementation of PPS, despite the shorter lengths of stay and reduced functional gain. Reasons for this improvement may be that patients are pleased with a shorter length of stay, or it may be the result of several quality-improvement initiatives that facilities undertook to counter the anticipated effects of PPS. One facility (rehabilitation hospital B) implemented a discharge-planning video to help patients prepare for discharge. This may have influenced the effect of lowered length of stay. In fact, all facilities in this study implemented additional discharge planning and trained staff on how to communicate functional goals with patients. Several facility-level improvements also occurred, such as expansions and additions of private rooms. It is difficult to isolate the effect of PPS on patient satisfaction, given the concurrent efforts at improving patient satisfaction. These results, therefore, must be interpreted in light of other studies that have shown an effect of decreased length of stay and decreased functional gain. Quality-improvement initiatives must be studied in isolation to see whether they actually improve satisfaction. Because events in real life do not occur in isolation, one must be prepared to combine results of various studies to truly determine the effect of an ongoing change as large as PPS.

Patient satisfaction is an important outcome from both a business perspective and as an indicator of quality care. Many hospitals track patient satisfaction to monitor quality-improvement activities. Owners of multiple hospitals compare patient satisfaction across facilities and nationally. Clinical and demographic characteristics vary widely across facilities, reflecting their affiliations and consequent referral patterns, geographic locations, and resources. Raw-score analysis does not take into account changing patient characteristics over time within a facility. Rehabilitation hospital A is a ter-

<table>
<thead>
<tr>
<th>Variables</th>
<th>Odds Ratio (95% CI)</th>
<th>SE</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rehabilitation hospital B*</td>
<td>1.65 (1.39–1.96)</td>
<td>0.15</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Rehabilitation hospital C*</td>
<td>3.50 (2.62–4.66)</td>
<td>0.51</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Rehabilitation hospital D</td>
<td>1.00 (0.80–1.26)</td>
<td>0.12</td>
<td>0.99</td>
</tr>
<tr>
<td>Motor FIM gain/12*</td>
<td>1.13 (1.04–1.24)</td>
<td>0.05</td>
<td>0.01</td>
</tr>
<tr>
<td>Cognitive FIM gain/5</td>
<td>1.09 (0.97–1.22)</td>
<td>0.06</td>
<td>0.14</td>
</tr>
<tr>
<td>Age</td>
<td>1.00 (1.00–1.01)</td>
<td>0.00</td>
<td>0.59</td>
</tr>
<tr>
<td>Gender</td>
<td>0.96 (0.83–1.12)</td>
<td>0.74</td>
<td>0.61</td>
</tr>
<tr>
<td>Respondent*</td>
<td>1.27 (1.07–1.50)</td>
<td>0.11</td>
<td>0.01</td>
</tr>
<tr>
<td>Discharge to home*</td>
<td>1.65 (1.31–2.07)</td>
<td>0.19</td>
<td>&lt;0.01</td>
</tr>
</tbody>
</table>

* Boldface, significant with P < 0.05.

<table>
<thead>
<tr>
<th>TABLE 3 Pre– and post–Prospective Payment System (PPS) observed and predicted excellent satisfaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facility</td>
</tr>
<tr>
<td>------------------------------------------</td>
</tr>
<tr>
<td>Rehabilitation hospital A</td>
</tr>
<tr>
<td>Rehabilitation hospital B</td>
</tr>
<tr>
<td>Rehabilitation hospital C</td>
</tr>
<tr>
<td>Rehabilitation hospital D</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

* Pre-PPS observed significantly different than post-PPS observed with P < 0.01.
† Post-PPS predicted significantly different than post-PPS observed with P < 0.01.
Boldface, statistically different from the reference value, which is post-PPS observed.
tiary free-standing urban hospital that cares for patients with more severe disabilities, whereas rehabilitation hospital C, a rural rehabilitation unit, had the highest admission FIM scores (data not shown). Respondent types may also vary depending on primary impairment. Patients with more severe disabilities may not be able to provide self-reports and are more likely to have a proxy respondent. The free-standing rehabilitation hospital (rehabilitation hospital A) had the lowest self-respondent percentage, and the trauma center and the rural rehabilitation hospital (rehabilitation hospitals D and C) had the highest proportion of self-respondents, consistent with the patients' functional status at these facilities. In this case, the predicted scores did not differ greatly from the observed scores. Even if risk adjustment does not significantly change satisfaction comparisons, it may allow practitioners to accept the results and minimize arguments that their patients are different.

There is no industry consensus on how to measure patient satisfaction or on which factors should be included in risk adjustment of patient satisfaction. The most significant variables in the model were the facility coefficients themselves, which revealed that facilities vary in patient satisfaction. Other variables such as age, gender, and respondent type were not significantly related to satisfaction. It may be that controlling for these characteristics is not necessary for an unbiased comparison of patient satisfaction. However, motor FIM gains and home discharge were significantly related to satisfaction. These findings suggest that patients are more satisfied if they have increased functional gain and are able to go home rather than to another post–acute care facility. The issue of whether functional gains affect satisfaction is debatable. As discussed in the introduction, many studies have refuted the point that motor gain increases satisfaction. In addition, the exclusion of length of stay as an independent variable lets us identify the effect of length of stay on satisfaction rather than controlling for the differences.

Because of the small changes in demographic characteristics and the small effects of the variables, however, the predicted scores did not vary greatly from the raw scores. Although the differences in patient characteristics account for some difference in reported satisfaction, they explain very little of the variance. However, one can imagine that if the difference in characteristics were very large, this could cause satisfaction scores to be different. The Centers for Medicare & Medicaid Services' 75% rule, which states that 75% of patients in rehabilitation facilities must have one of 13 diagnostic categories, may alter rehabilitation hospitals' patient diagnostic and sociodemographic characteristics to such a degree such risk adjustment could significantly affect patient-satisfaction scores.

A major limitation of the study is the difficulty of measuring satisfaction reliably. As mentioned before, satisfaction is a complex construct, and it is difficult to assess. A single item was chosen, and although this item correlated well with other items, we cannot estimate the reliability of a single item. Our ability to generalize the results is limited by the small sample of facilities. Furthermore, all facilities studied share a similar corporate affiliation, even though they are in different communities. The rehabilitation hospitals studied share a similar organizational philosophy and are guided by the same values. Care should be taken when generalizing these results to all rehabilitation hospitals nationally. These differences may not reflect a significant difference in satisfaction for purposes of consumer loyalty or for measures of healthcare quality.

Further research would involve using a larger sample of patients to determine the effect of PPS at many different rehabilitation hospitals. The factors that affect satisfaction also need to be clarified. Length of stay, functional gains, and other demographic factors showed a small effect on satisfaction, and larger samples would help clarify this issue. The issue of isolating the effects of a quality-improvement initiative or a new payment system such as PPS can be resolved by measuring specific aspects of satisfaction such as satisfaction with discharge or satisfaction with teaching. Looking at these variables specifically as dependent variables would require further delineation of specific independent variables.

CONCLUSION

Despite a decrease in motor FIM gains after PPS, patient satisfaction increased after PPS implementation. Risk adjustment for these characteristics had a small effect on overall satisfaction. The risk of bias in estimating overall satisfaction is such that respondent type, discharge to home, and functional status should be included in models that compare satisfaction across facilities.

ACKNOWLEDGMENTS

We would like to acknowledge insightful comments and helpful suggestions from Ming Zhang, PhD, Trudy Mallinson, PhD, Anne Deutsch, RN, PhD, Deborah Dobrez, PhD, and Elizabeth Durkin, PhD. Funding for this project was provided by the National Institute on Disability and Rehabilitation Research for a project titled "Health Services Research DRRP on Medical Rehabilitation" (H133A030807).

REFERENCES


11. *Envisioning the National Health Care Quality Report*. Washington, DC, Institute of Medicine, 2001


29. Carter GM, Paddock SM: Preliminary Analyses of Coding and Case Mix Under the Inpatient Rehabilitation Facility Prospective Payment System. Santa Monica, Ca, RAND Health, 2005


32. Medicare Program; Changes to the Criteria for Being Classified as an Inpatient Rehabilitation Facility. *Federal Register* 2004; 69(89):25752
ADL Assessment of Nondisabled Turkish Children with the WeeFIM Instrument

ABSTRACT

Objective: To evaluate the WeeFIM instrument’s reliability and internal construct validity for the Turkish child population.

Design: License was taken from UDSmr to use the WeeFIM instrument. For the reliability and validity studies of the Turkish translation of the WeeFIM instrument, 573 Turkish nondisabled children were included in the study. The reliability of the instrument was assessed by Cronbach’s alpha coefficient, intraclass correlation coefficient (ICC), and test–retest reliability. Internal construct validity was assessed by both using Rasch unidimensional measurement model and testing for differential item functioning for age and gender.

Results: Cronbach $\alpha$ value was 0.99 for motor WeeFIM rating and 0.99 for cognitive WeeFIM rating. ICC was 0.81 for motor WeeFIM rating and 0.92 for cognitive WeeFIM rating. The internal construct validity of the Turkish translation of the WeeFIM instrument was confirmed by excellent fit to the Rasch measurement model. Two subscales were found from the principal component analysis of standardized residual correlation for items. Among the items, bowel management, bladder management, eating, and comprehension showed considerable levels of misfit.

Conclusions: The Turkish translation of the WeeFIM instrument is valid, reliable, and practical for the Turkish child population. Further studies are required to determine the cross-cultural validity of the instrument.

Key Words: WeeFIM, Functional Assessment, Rasch Analysis, Validity, Reliability
Although pediatric evaluation methods are not used as broadly as the methods applied to assess adult rehabilitation, the number of recent studies has increased. One of the most commonly used methods for pediatric functional evaluation is the WeeFIM instrument, a measure of functional independence in children. It was developed in 1993, arising from the FIM instrument designed to measure severity of disability in adults. Recent studies have demonstrated the reliability and validity of the WeeFIM instrument for both disabled and nondisabled children and have emphasized that it is an excellent evaluation method.

The normative samples of the WeeFIM instrument for nondisabled American and Japanese children have been obtained by the studies carried out before. Studies performed for different populations have emphasized that age, sociocultural and economic status of families, and environmental conditions should be taken into consideration when evaluating the independence level of activities of daily living (ADL) of children.

Although ADL and handicap measurements with proven validation and reliability are available for Turkish adults, there is not any ADL measurement method that is valid and reliable for Turkish children. Therefore, we first used the Turkish translation of the WeeFIM instrument in 45 children with cerebral palsy (CP) and 41 Turkish nondisabled children as a preliminary study. We discovered that all WeeFIM ratings of the small group of Turkish nondisabled children examined in the previous study did not reflect normative WeeFIM ratings of the Turkish child population. Therefore, we decided to conduct a study with a larger case group to prove the validity and reliability of the WeeFIM instrument in our population. Our purpose was to evaluate the WeeFIM instrument’s reliability and internal construct validity for the Turkish child population.

**METHODS**

**Subjects**

For the reliability and validity studies, 573 Turkish nondisabled children (295 boys and 278 girls) aged 7–92 mos without developmental delays were included in the study. The mean age of the children was 39.1 mos (SD 23.9, median 38 mos). The cases were obtained from two health centers available in different districts of Ankara. All of them were nondisabled and attending to these centers for general health checks and vaccination purposes. The purposes and procedures of the study were explained, and informed consent was obtained from the parents.

**The WeeFIM Instrument**

The WeeFIM instrument contains 18 measurement items, divided into six areas: self-care (eating, grooming, bathing, dressing upper body, dressing lower body, toileting), sphincter control (bowel management, bladder management), mobility (chair/bed/wheelchair transfer, toilet transfer, tub transfer), locomotion (crawling/walking/wheelchair, stair climbing), communication (comprehension, expression), and social cognition (social interaction, problem solving, memory). A seven-level ordinal scoring system ranging from 7 (complete independence) to 1 (total assistance) is used to rate performance. The minimum possible total score is 18 (total dependence in all skills); the maximum possible score is 126 (complete independence in all skills). Because our sample consisted of nondisabled children, we assessed the walking task taking into consideration both crawling and walking, not wheelchair use.

Permission was obtained from the Uniform Data System for Medical Rehabilitation (UDSmr) to use the WeeFIM instrument. The WeeFIM II System Interim Clinical Guide 2003 book was obtained from UDsmr.

**Translation Procedure of the WeeFIM Instrument**

The original WeeFIM form was translated into Turkish independently by a physiatrist with advanced English and an English teaching professor who had lived in the United States. These Turkish translations of the WeeFIM instrument were then translated back into English by another team that consisted of a professional technical translator and a physiatrist who was fluent in English. These texts were compared with each other, and the final Turkish translation of the WeeFIM instrument was formed by a team consisting of four physiatrists and other relevant healthcare professionals specializing in the field of pediatric rehabilitation.

In the study, one physiatrist performed the WeeFIM instrument by combining direct observation of the children, with the help of the interviews with primary caregivers consisting of 95.6% mothers, 2.6% fathers, and 1.7% other relatives. After 5 days, 102 children (55 girls and 47 boys) aged 7–89 mos were assessed again by the same physiatrist.

**Statistical Analysis**

The scientific quality of the WeeFIM instrument was determined through a range of analysis, which includes tests for reliability and validity. The motor and cognitive items of the WeeFIM instrument were analyzed both together and separately.
Assessment of Reliability

Three types of reliability were assessed: internal consistency, intraclass correlation coefficient (ICC), and the test–retest reliability. One way of assessing reliability is the consistency or repeatability of the measures; this method depends on how much of the variation in measures is attributable to random or chance errors. Internal consistency of the Turkish translation of the WeeFIM instrument was provided by Cronbach’s α coefficient, which is defined as the interrelationship of items within a scale. ICC (one-way random-effect model) was used as a second way of assessing reliability. ICCs are often used as reliability coefficients in evaluations of items that are deemed to be in the same category or class. They are the ratios of between-rating variance to total variance. Although the ICCs are defined in terms of proportions of variance, it is possible for empirical estimates to be negative; the estimates all have upper bounds of 1, but no lower bounds. The test–retest reliability was assessed by Spearman’s correlation coefficient on 102 nondisabled children.

Assessment of Internal Construct Validity

Internal construct validity was assessed using the Rasch unidimensional measurement model. The Rasch model is a unidimensional model that asserts that the easier the item (task), the more likely it will be passed, and the more able the person, the more likely they will pass an item (or be able to do a task) compared with a less able person. The Rasch model is now used extensively within the medical outcomes field to test the unidimensionality of scales, the ways in which their categories function, and whether items are biased for key subgroups such as age or gender (differential item functioning [DIF]).

The Rasch measurement model transforms ordinal raw scores into interval measures that are objective, fundamental, and linear. Rasch measures provide more information than the observed raw scores because the former provide sample-free item difficulties and test-free person measures. The Rasch measurement model can also identify items that are redundant and those that do not fit the model. Two indicators of fit are mean square OUTFIT and mean square INFIT statistics. These statistics are measures of how well a WeeFIM item fits to the Rasch model. The reasonable item mean square range for INFIT and OUTFIT is from 0.6–1.4.

Another important aspect of the internal integrity of a scale is the absence of item bias or DIF. At a given level of disability, it is important that the response to any item is unaffected by group membership. In the current study, DIF was examined by age and gender. Bonferroni corrections were applied to both fit and DIF statistics because of the number of tests undertaken. Because of multiple significance tests, P values less than 0.008 were considered statistically significant.

Differences among four age groups (6–21, 22–45, 46–62, and 63–100 mos) according to Msall et al. for each item were evaluated by Kruskal–Wallis variance analysis. When the P value from the Kruskal–Wallis test statistics was statistically significant, a multiple comparison test was used to determine which groups differed from which others.

Data were analyzed by using SPSS for Windows 11.5 and the Rasch model computer program WINSTEPS.

RESULTS

Median WeeFIM ratings and distribution of gender of the subjects for each age group are presented in Table 1.

Reliability

Cronbach α value was 0.99 for motor WeeFIM ratings and 0.99 for cognitive WeeFIM ratings. The ICC (one-way random-effect model) was 0.81 for motor WeeFIM ratings and 0.92 for cognitive WeeFIM ratings. Test–retest reliability was good, with high correlations between the two time points (r = 0.996, P < 0.001 for total; r = 0.954, P < 0.001 for motor; r = 0.987, P < 0.001 for cognitive).

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Characteristics of the subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age Range, mos</td>
<td>Sex</td>
</tr>
<tr>
<td>7–21</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>F</td>
</tr>
<tr>
<td>22–45</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>F</td>
</tr>
<tr>
<td>46–62</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>F</td>
</tr>
<tr>
<td>63–92</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>F</td>
</tr>
<tr>
<td>Total</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>F</td>
</tr>
</tbody>
</table>

M, male; F, female; min, minimum; max, maximum.
Internal Construct Validity
Internal Construct Validity of the WeeFIM instrument by Rasch model

The internal construct validity of the Turkish translation of the WeeFIM instrument was confirmed by its excellent fit to the Rasch measurement model. We found two subscales from the principal-component analysis of standardized residual correlation for items. Although the stair climbing, eating, grooming, walking, and bathing items did not load into any subscale, the remaining motor items loaded into the first subscale, and all cognitive items loaded into the second subscale. Therefore, we performed Rasch analysis on the WeeFIM instrument and analyzed fit statistics with all 18 items as combined and motor and cognitive items as subdivided. Table 2 shows both combined and separate analyses performed on all 573 nondisabled children.

For the motor items, bathing, lower dressing, upper dressing, and toileting were difficult items, whereas walking/crawling and stair climbing were the easiest. The difficulty patterns for the four age groups were almost identical, except for bowel management and bladder management (Fig. 1). For the cognitive items, memory and problem solving were the more difficult items, whereas comprehension and expression were easier. The difficulty patterns for the four age groups were almost identical (Fig. 2). Among the items, bowel management, bladder management, eating, and comprehension show considerable levels of misfit, as determined by both INFIT and OUTFIT statistics (Table 2).

Internal Construct Validity: DIF

There were statistically significant differences among the age groups in terms of motor items. The first age group (7–21 mos) was significantly different from the other age groups in terms of eating, bladder and bowel management, and grooming. The age groups did not present any statistically significant differences among themselves with respect to toileting, dressing the lower body, bed/chair transfer, stair climbing, or walking. For motor items, there was not any statistically significant difference among age groups in terms of gender.

There were statistically significant differences among the age groups in terms of cognitive items. The first age group was significantly different from the second (22–45 mos), and the third age group (46–62 mos) was significantly different from all age groups in terms of comprehension. With respect to social interaction, problem solving, and memory items, there were no significant differences among the age groups. Regarding cognitive items, there were no statistically significant differences among age groups in terms of gender.

| TABLE 2 Item difficulties (by logits) and fit statistics for combined and divided items |
|---|---|---|---|---|---|
| | Item Difficulties (Error) | INFIT (MNSQ) | OUTFIT (MNSQ) |
| | Combined | Divided | Combined | Divided | Combined | Divided |
| Motor Items |
| Eating | 0.79 (0.06) | 0.99 (0.06) | 1.32 | 1.48 | 7.21 | 9.90 |
| Grooming | 0.83 (0.06) | 1.03 (0.06) | 0.70 | 0.72 | 0.80 | 0.99 |
| Bathing | 1.98 (0.05) | 2.17 (0.05) | 0.81 | 0.75 | 0.90 | 0.85 |
| Upper dressing | 1.02 (0.06) | 1.22 (0.06) | 0.73 | 0.60 | 0.57 | 0.51 |
| Lower dressing | 1.09 (0.06) | 1.29 (0.05) | 0.61 | 0.48 | 0.50 | 0.42 |
| Toileting | 0.95 (0.06) | 1.15 (0.06) | 0.93 | 0.73 | 0.74 | 0.60 |
| Bladder management | −0.20 (0.06) | −0.07 (0.07) | 1.63 | 1.52 | 0.97 | 0.87 |
| Bowel management | −0.36 (0.06) | −0.24 (0.07) | 1.83 | 1.72 | 0.99 | 0.86 |
| Bed/chair transfer | −0.72 (0.07) | −0.65 (0.07) | 0.79 | 0.62 | 0.49 | 0.43 |
| Toilet transfer | −0.45 (0.06) | −0.34 (0.07) | 0.78 | 0.60 | 0.50 | 0.39 |
| Tub/shower transfer | −0.19 (0.06) | −0.05 (0.07) | 0.86 | 0.75 | 0.59 | 0.54 |
| Walk or wheelchair | −3.44 (0.09) | −3.73 (0.09) | 1.05 | 0.86 | 0.31 | 0.24 |
| Stairs | −2.54 (0.08) | −2.76 (0.09) | 0.83 | 0.90 | 0.38 | 0.53 |
| Cognitive Items |
| Comprehension | −0.95 (0.07) | −2.97 | 1.41 | 1.23 | 1.66 | 1.16 |
| Expression | 0.08 (0.06) | −0.56 | 0.91 | 0.94 | 0.90 | 0.96 |
| Social interaction | 0.12 (0.06) | −0.48 | 0.83 | 0.86 | 0.93 | 0.80 |
| Problem solving | 0.92 (0.06) | 1.78 | 0.73 | 0.87 | 0.72 | 0.80 |
| Memory | 1.07 (0.06) | 2.22 | 0.64 | 0.81 | 0.59 | 0.71 |

MNSQ, mean square.
DISCUSSION
Msall et al.\textsuperscript{28} showed that the WeeFIM instrument was a reliable and valid measurement for the functional assessment of children aged between 6 mos and 7 yrs with CP, Down syndrome, spina bifida, extreme prematurity, and limb deficiency. They also revealed that it is a reliable and valid measurement method for children aged between 6 mos and 12 yrs and that it can be used both during preschool and school years. Recent studies have demonstrated the reliability and validity of the WeeFIM instrument for both disabled and nondisabled children.\textsuperscript{2,3} However, the results of the functional measurement scales may show variations in different countries because of the differences in sociocultural behaviors, including feeding, bathing, grooming, and toilet habits. To our knowledge, the literature does not contain any studies about the WeeFIM instrument demonstrating the normative data of the Turkish population. One of our aims in this study was to fill this gap in the literature.

Heinemann et al.\textsuperscript{29} and Granger et al.\textsuperscript{30} analyzed the scale quality of the FIM instrument using Rasch analysis. Tsuji et al.\textsuperscript{6} used Rasch rating scale analysis (RSA) to examine WeeFIM item difficulty. We have also employed Rasch analysis in our study to establish the internal construct validity of the WeeFIM instrument for Turkish children.

In this study, we found two subscales from the principal-component analyses of standardized residual correlation for items. Although all motor items loaded into the first subscale, all cognitive items loaded into the other. Therefore, we performed Rasch analysis on the WeeFIM instrument and analyzed fit statistics with all 18 items as combined and motor and cognitive items as subdivided. Chen et al.,\textsuperscript{31} in a retrospective study on

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure1}
\caption{WeeFIM item difficulties for motor items (by logits) for nondisabled Turkish children aged 7–92 mos.}
\end{figure}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure2}
\caption{WeeFIM item difficulties for cognitive items (by logits) for nondisabled Turkish children aged 7–92 mos.}
\end{figure}
814 children with disability, applied Rasch RSA to the data obtained with the WeeFIM instrument. They also identified two dimensions (i.e., motor and cognitive) by principal-component analysis of the item residuals. Tsuji et al. in a study they performed with the WeeFIM instrument on 225 children without any developmental delay, observed that when the WeeFIM items were divided into two groups of motor and cognitive items to minimize misfit, the degree of misfit was acceptable. Based on these findings, it can be said that the WeeFIM instrument has two dimensions as motor and cognitive.

In the study, the Turkish translation of the WeeFIM instrument’s reliability was good, with excellent internal consistency and test–retest reliability. Internal construct validity was confirmed by good fit to the Rasch measurement model. We have observed that for the motor items, bathing, lower dressing, upper dressing, and toileting were difficult items, whereas walking and stair climbing were the easiest ones. For the cognitive items, memory and problem solving were more difficult items, whereas comprehension and expression were easier. Our findings are comparable with the results found with nondisabled Japanese children. Tsuji et al. concluded that stair climbing, bed/chair transfer, and walking were the easiest items, whereas grooming, bathing, and bladder control were the hardest items for nondisabled Japanese children. Chen et al., dividing 814 disabled children into three age groups (<3, 3–7, and >7 yrs of age) in their study, reported that the order of motor-item difficulty varied across age groups.

The fit statistics conducted in the present study are acceptable, except for eating, bladder and bowel management, and comprehension. Comprehension seemed to be misfit in the analysis of all 18 items, although it did not seem so in motor and cognitive assessments performed separately. Similarly, Tsuji et al. also found that the comprehension item was a misfit. We can attribute the misfit status of the eating item by the fact that children’s diets become richer as they grow up, which is a more distinct situation, especially in our society. Therefore, it may be more difficult for the concerned child to perform eating functions properly, promptly, and reliably using a fork and knife set. On the other hand, if an 18-mo-old baby who has been fed solely on the bottle is able to take his or her bottle alone and self-feed reliably, she or he may receive a higher-than-expected rating for this task. In addition, another reason may be the incapability of the WeeFIM instrument to investigate, in detail, food types and how the functions are performed. Wong et al., in a study performed on 445 nondisabled Chinese children, have found a significant correlation between the eating rating of the WeeFIM instrument and the achievement age of using chopsticks. They mention that chopsticks use would make the eating task more difficult. Misfits of this type may be attributed to different sociocultural contexts.

Misfit status of bladder and bowel management may be attributable to the fact that these functions are acquired at the age of 2–3 yrs in many children, and the associated ratings may remain unchanged despite growing age. In the preliminary study we conducted on children with CP and nondisabled children, we found that all WeeFIM ratings except for sphincter control showed increases with age in nondisabled children. Bowel and bladder management was also reported as a misfit item on the adult FIM instrument. Furthermore, as emphasized in earlier studies, this situation may occur because both physiologic and performance (both autonomic and voluntary neurological control mechanisms for bowel and, especially, bladder management) aspects of these functions are rated.

The comparison of item difficulty among the four age groups in this study revealed differences in difficulty levels among some age groups, especially with regard to eating and bladder and bowel management. These differences may be related to children’s ability and task complexity. For example, in contrast to the other items, the eating item seems easier in the first age group (Fig. 1). This might be explained by the comfort provided by the use of feeding bottles in this age group. Chen et al. reported that only eating and stair climbing items showed misfit in their study conducted on school-age disabled children. They also stated that toileting was the most difficult for toddlers but was relatively easier for older children, walking was easiest for toddlers but was relatively harder for older children, and grooming was difficult for toddlers but was easier for older children. In contrast, the difficulty levels revealed by our study for toileting, walking, and transfer activities were similar in all age groups because our sample consisted of nondisabled children in contrast with Chen’s study. Squatting toilets are still common, especially in lower socioeconomic groups in Turkey. However, closet-type lavatories are becoming more common in schools and residential dwellings. Therefore, we found no differences among the age groups regarding the difficulty levels for toilet transfer.

**CONCLUSION**

Rasch analysis of the WeeFIM instrument has been shown to incorporate two dimensions, one for the motor items and the other for the cognitive items. The Turkish translation of the WeeFIM instrument is valid, reliable, and practical for the
Turkish child population. The WeeFIM instrument is convenient for use in pediatric rehabilitation. Further studies are required to determine the cross-cultural validity of the Turkish translation of the WeeFIM instrument.

ACKNOWLEDGMENTS

The authors would like to extend their appreciation to the Uniform Data System for Medical Rehabilitation, Center for Functional Assessment Research at the State University of New York at Buffalo. The use of the WeeFIM instrument to collect data for this research study was authorized and conducted in accordance with the terms of a special purpose license granted to the Licensee by Uniform Data System for Medical Rehabilitation, a division of UB Foundation Activities, Inc. (UDSmr). The patient data collected during the course of this research study have been or will be subjected to UDSmr’s standard data processing procedures or are otherwise comparable with data processed by UDSmr.

REFERENCES

1. UDSMR. Guide for the Uniform Data Set for Medical Rehabilitation for Children (WeeFIM), version 4.0. Buffalo, NY: State University of New York at Buffalo, 1993
Authors:
Ta-Shen Kuan, MD, MS
Yueh-Ling Hsieh, PhD, RPT
Shu-Min Chen, MD
Jo-Tong Chen, MD, PhD
Wei-Chang Yen, MD
Chang-Zern Hong, MD

Affiliations:
From the Department of Physical Medicine and Rehabilitation, College of Medicine, National Cheng Kung University, Tainan, Taiwan (T-SK, S-MC, J-TC, W-CY, C-ZH); Department of Physical Therapy, Hungkuang University, Sha-Lu, Taichung, Taiwan (Y-LH, C-ZH); and Department of Rehabilitation Medicine, Han-Ming Hospital, Chang-Hua, Taiwan (C-ZH).

Correspondence:
All correspondence and requests for reprints should be addressed to Chang-Zern Hong, MD, Department of Physical Therapy, Hungkuang University, 34 Chung-Chie Road, Sha-Lu, Taichung 433, Taiwan.

Disclosures:
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 authors or upon any organization with which the authors are associated.

ABSTRACT

Objective: This study was designed to investigate the correlation between the irritability of the myofascial trigger point (MTrP) and the prevalence of endplate noise (EPN) in the MTrP region of human skeletal muscle.

Design: Twenty normal subjects with latent MTrPs and 12 patients with active MTrPs in the upper trapezius muscles were recruited for this study. The patients reported the subjective pain intensity of the active MTrP (0–10). The MTrP and an adjacent non-MTrP site were confirmed and marked for the measurement of pressure pain threshold (with a pressure algometer) and the prevalence of EPN (with electromyographic recordings).

Results: The prevalence of EPN in the MTrP regions was significantly higher ($P < 0.01$) in the active MTrPs than in the latent ones. However, no EPN could be found in the non-MTrP region near either the active or the latent MTrPs. The pain intensity and the pressure pain threshold were highly correlated with the prevalence of EPN in the MTrP region ($r = 0.742$ and $-0.716$, respectively).

Conclusions: The irritability of an MTrP is highly correlated with the prevalence of EPN in the MTrP region of the upper trapezius muscle. The assessment of EPN prevalence in an MTrP region may be applied to evaluate the irritability of that MTrP.

Key Words: Electromyography, Endplate Noise, Muscle, Trigger Point, Pain
Myofascial trigger point (MTrP) presence is characteristic of myofascial pain syndrome. An MTrP has been defined as a highly localized painful or sensitive spot in a palpable taut band of skeletal muscle fibers. An active MTrP is an irritable spot with spontaneous pain or pain in response to movement, and a latent MTrP is a tender spot with pain or discomfort in response to compression only. Two important characteristics of MTrPs are referred pain and a local twitch response (LTR). LTRs can be elicited by snapping palpation on the MTrP in some muscles, or by needling in almost all cases. More LTRs can be elicited by needling of a highly irritable MTrP than a less painful one.

The mechanism of LTR has been extensively studied on the animal model developed by Hong and Torigoe. The most tender spot in the hamstring muscle of rabbit could be identified by observing the animal’s behavior. This hyperirritable spot was marked and the animal was anesthetized. It was found that many LTRs (similar to those observed in humans) could be elicited by needling of this hyperirritable spot, but this was possible in very few or none in the control sites. It was also found that LTRs could not be elicited after the transection of the innervation nerve (sciatic nerve). LTRs could also not be elicited in the hamstring muscles immediately after transection of the upper thoracic spinal cord, but they recovered almost to the original level after the spinal shock period (about 2.5 hrs later). It has been concluded that LTR is elicited via the spinal reflex by stimulating the sensitive site in the MTrP region. Hong and Simons have hypothesized that there are multiple sensitive loci in the MTrP region. Mechanical stimulation of this tiny locus can elicit pain, referred pain, and LTR (with high-pressure stimulation, such as needling). This tiny sensitive spot has been defined as a sensitive locus, or an LTR locus.

In 1993, Hubbard and Berkoff recorded spontaneous electrical activity in the MTrP region of the upper trapezius muscle. Similar electrical activity could also be recorded from the rabbit skeletal muscles. Hubbard and Berkoff described this activity as the action potential generated from the intrinsic muscle fibers when a muscle spindle is mechanically irritated. However, this low-amplitude electrical activity is similar to the endplate noise (EPN) rather than action potentials. After reviewing old literature, Simons concluded that EPN is a consequence of excessive acetylcholine leakage, which may cause focal contracture in the endplate zone to form the contracture nodule or taut band. In recent studies, Simons et al. have demonstrated that there is a much higher prevalence of EPN in the MTrP region compared with the non-MTrP site in either animals or human subjects.

It has been suggested that the amount of sensitive loci (LTR loci) in an MTrP region is proportionate to the degree of irritability of the MTrP. However, it is unclear whether the amount of active loci (EPN loci) is also related to the degree of irritability of MTrP. This study is designed to investigate the correlation between the prevalence of EPN loci and the degree of irritability (pain intensity, pain threshold) in the MTrP region.

### MATERIALS AND METHODS

#### General Design

Normal subjects with latent MTrPs and patients with active MTrPs in one side of the upper trapezius muscles were included in this study. Patients were initially requested to report the subjective pain intensity of the active MTrP according to a numeric rating scale (0–10). For each subject (either normal subject or patient), pressure pain threshold was measured at both MTrP and non-MTrP sites. Then, the prevalence of EPN was assessed with needle electromyographic (EMG) recordings in both the MTrP region and non-MTrP region of the upper trapezius muscle for each subject.

#### Subjects

Twenty normal subjects (age 43.5 ± 11.1) with latent MTrPs and 12 patients (age 43.7 ± 11.9) with active MTrPs in the upper trapezius muscle were recruited for this study. Subjects in both groups were matched for age and sex. They all signed the informed consents as approved by the human subject research committee of the National Cheng Kung University, Tainan, Taiwan. The general criteria for selection of subjects included (1) no acute or serious medical problems, (2) no neurologic disorders other than muscle pain, (3) no coagulopathy or any other bleeding disorder, (4) no serum hepatitis B or acquired immunodeficiency syndrome, (5) no pain therapy (pain medications, physical therapy, etc.) for at least 2 wks, and (6) no cognitive impairment or psychiatric disorder.

#### Identification of MTrPs and Non-MTrPs

##### Normal Subjects with Latent MTrPs

Normal subjects who had latent MTrPs in the upper trapezius muscles in either side or both sides were recruited. If MTrPs existed in both sides of upper trapezius muscles, only one side was randomly selected for the study. The latent MTrP was identified on the basis of the following criteria: (1) a localized tender spot (without spontaneous pain) in a palpable taut band of muscle fibers, and (2) characteristic and consistent referred pain when it...
was compressed firmly. As soon as this latent MTrP was confirmed, it was marked by an investigator who did not perform the assessment of pain threshold or EPN prevalence.

Patients with Active MTrPs

Patients who had active MTrPs in the upper trapezius muscles in either side or both sides were recruited. If MTrPs existed in both sides of upper trapezius muscles, only one side was randomly selected for this study. The active MTrP was identified on the basis of the following criteria, as recommended by Simons: (1) a localized tender spot in a palpable taut band of muscle fibers, (2) recognized pain (as the usual clinical complaint) when the tender spot was compressed, and (3) characteristic and consistent referred pain. Similarly, this active MTrP was marked by the same investigator who marked the latent MTrP.

Localization of Non-MTrPs

After identification of the MTrP site in each subject, a control site (non-MTrP site) some distance away from the MTrP site was also marked by the same investigator who marked the MTrP site.

Subjective Pain Intensity (Numeric Rating Scale)

Initially, the patient was requested to report pain intensity of the selected MTrP on the basis of the numeric rating scale. The patient was informed of the numeric rating scale from 0 to 10. Zero represents no pain at rest or during movement, and 10 represents the worst pain ever experienced in one’s life. The other numbers represent different degrees of pain level. In general, pain intensity below 5 is considered a tolerable level.

Pressure Pain Threshold Measurement

Measurement of pressure pain threshold on the MTrP and non-MTrP sites was performed before the EMG recordings. A pressure algometer was used for this measurement. A study by Ohrbach and Gale found that pain threshold measurement with this algometer was reliable and had no significant differences among multiple trials in painful muscles. Initially, the algometer was placed on the marked MTrP site, perpendicular to the surface of the skin. The pressure of compression was increased gradually at a rate of movement approximately 1–2 mm/sec until the subject began to feel any pain or discomfort (for latent MTrPs) or until the patient began to feel an increase of pain or discomfort (for active MTrPs). At that point, the subject informed the examiner by saying “yes.” The compression stopped as soon as the subject said “yes.” The subject was asked to remember this level of pain discomfort and to apply the same criterion for the consecutive measurements. The reading of pressure (kg/cm²) at that point was considered the pressure pain threshold level. The subject might demonstrate pain by pulling away or grimacing, which would indicate that the pain threshold had been exceeded. If this was the case, the subject would be given instructions again, and a repeat measurement would be taken to ensure that the “real” threshold was obtained. Three repetitive measurements at an interval of 30–60 secs were performed at each site. The average values of the three readings were used for data analysis. The same routine was also performed at the control site (non-MTrP site).

Two investigators who participated in measuring the threshold were trained extensively to minimize the intratester and intertester variability. They were blinded as to whether subjects had latent or active MTrPs.

Electrophysiological Recordings

Equipment

A four-channel Nicolet EMG machine was used to record the electrical activity from MTrPs by using disposable monopolar Teflon-coated EMG needle electrodes.

Settings

As shown in Figure 1, the first channel recorded the electrical activity from the active electrode (experimental needle electrode) in the MTrP region. The reference needle electrode was placed in a site approximately 3–4 cm from MTrP. The second channel recorded the electrical activity from the control site. For the control recordings, the active recording electrode was placed in a non-MTrP region outside the endplate zone. The reference electrode was connected to the reference electrode of the first channel through a bridge connector (to form a common reference electrode). A ground electrode was placed on the skin near the recording sites. For both channels, the sensitivity of recording was set at 20 mV per vertical division, and the time-sweep speed was set at 10 msecs per horizontal division. The high-frequency filter was set at 10 kHz, and the low-frequency filter was set at 20 Hz.

Procedure

Regarding the assessment of EPN, the prevalence of EPN in 24 sites of three needle-insertion tracks in the MTrP region was measured as described in the previous reports (Fig. 2). In this way, only one penetration into the skin was required. We did not measure the amplitude and duration of the EPN because the amplitude and duration varied during needle movement either because of changes in electrode distance or direct
mechanical injuries to the muscle fibers. In the experimental study of one MTrP site, the exploring electrode (active recording needle electrode) was inserted progressively into the MTrP region. The needle was advanced gently and slowly through the least possible distance (usually 1–2 mm) by simultaneously rotating the needle to facilitate smooth entry through the tissue. After eight thrusts (advancements of the needle in one track), the needle was pulled out to the original insertion depth and reinserted in a slightly different direction (a near track). This procedure was repeated again to explore a total of 24 thrusts (eight thrusts per insertion; three insertions). All occurrences of EPN were recorded. As soon as an EPN appeared, the needle remained there without further movement. Sample EMG recordings (Fig. 1) were taken until the amplitude of the EPN activity became indistinguishable from that recorded from the control needle. The time required for the disappearance of EPN was less than 10 mins, usually 2–3 mins. Advancement of the needle continued after disappearance of the EPN. The subject reported the pain intensity and feeling when an EPN was recorded. The routine used at the control site (non-MTrP site) was the same as that used at the MTrP site with regard to insertion procedure. The investigator who performed this procedure was blinded as to whether subjects had latent or active MTrPs.

**Data Analysis**  
**Comparison Between Latent and Active MTrPs**

The data of pressure pain threshold and prevalence of EPN in both MTrP and non-MTrP region were analyzed for the differences between normal subjects with latent MTrPs and patients with active MTrPs according to a t test (pain threshold data) or a Fisher test (EPN data). A P value less than 0.01 was considered statistically significant.

**FIGURE 1** Schematic arrangement of electrodes and typical examples of recordings.

**FIGURE 2** Schematic of insertion procedure during the assessment of EPN prevalence.
Comparison Between MTrP and Non-MTrP Sites

The data of pressure pain threshold and prevalence of EPN in both normal subjects and patients were analyzed for the differences between MTrP region and non-MTrP region according to a paired t test (pain threshold data) or a Fisher test (EPN data). A P value less than 0.01 was considered statistically significant.

Correlation Between the Prevalence of EPN and the Irritability (Pain Intensity or Pressure Pain Threshold) in the MTrP Region

The data of the pain intensity and the average values of three readings for each threshold measurement in all subjects (both normal subjects and patients) were analyzed for their correlation with the prevalence of EPNs. Pearson linear regression analysis was used to analyze statistical significance.

RESULTS

Latent MTrP (Normal Subjects) vs. Active MTrP (Patients)

As shown in Table 1, age and sex were fairly matched for patients and normal subjects. All patients had pain in the MTrP regions, but none had pain in the non-MTrP regions (with a mean pain intensity of 0.0 ± 0.0). Every patient reported pain localized in the shoulder without referral. The mean pressure pain threshold, either at the MTrP site (1.9 ± 0.3 kg/cm²) or at the non-MTrP site (4.1 ± 0.6 kg/cm²), was significantly lower (P < 0.01) in the patients with active MTrPs compared with the normal subjects with latent MTrPs (3.7 ± 0.7 and 5.9 ± 1.0 kg/cm², respectively). The prevalence of EPN in the MTrP regions was also significantly higher (P < 0.01) in the active MTrPs (7.2 ± 4.2 per 24 sites) than in the latent MTrPs (2.8 ± 1.4 per 24 sites).

MTrP vs. Non-MTrP

As expected, the mean pressure pain threshold at the MTrP site (3.7 ± 0.7 or 1.9 ± 0.3 kg/cm²) was significantly lower (P < 0.01) than at the non-MTrP site (5.9 ± 1.0 or 4.1 ± 0.6 kg/cm²) for either latent or active MTrPs. However, no EPN could be recorded in the non-MTrP region for either latent or active MTrPs (Table 1).

Prevalence of EPN and Subjective Pain Intensity of MTrP

The high correlation (r = 0.742) between the prevalence of EPN and the pain intensity in the MTrP region for all subjects is demonstrated in Figure 3. This was also true when the normal subjects were excluded from the population for analysis.

Prevalence of EPN and Pressure Pain Threshold of MTrP

For all subjects, the prevalence of EPN was highly inversely correlated (r = −0.716) with the pressure pain threshold (Fig. 4). Similar to a previous study,¹⁴ the pain intensity and the pressure pain threshold are also strongly inversely correlated.

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Pain intensity, pressure pain threshold, and prevalence of endplate noise (EPN) in the myofascial trigger point (MTrP) region and non-MTrP region</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of subjects</td>
<td>All Subjects</td>
</tr>
<tr>
<td>Male</td>
<td>14</td>
</tr>
<tr>
<td>Female</td>
<td>18</td>
</tr>
<tr>
<td>Age</td>
<td>43.5 ± 11.2</td>
</tr>
<tr>
<td>Pain intensity (0–10)</td>
<td></td>
</tr>
<tr>
<td>MTrP site</td>
<td>2.3 ± 3.1</td>
</tr>
<tr>
<td>Non-MTrP site</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>Pressure pain threshold, kg/cm²</td>
<td></td>
</tr>
<tr>
<td>MTrP site</td>
<td>3.0 ± 1.0</td>
</tr>
<tr>
<td>Non-MTrP site</td>
<td>5.3 ± 1.2</td>
</tr>
<tr>
<td>Difference</td>
<td>P &lt; 0.01</td>
</tr>
<tr>
<td>Prevalence of EPN, no./24 sites</td>
<td></td>
</tr>
<tr>
<td>MTrP site</td>
<td>4.6 ± 3.6</td>
</tr>
<tr>
<td>Non-MTrP site</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>Difference</td>
<td>P &lt; 0.01</td>
</tr>
</tbody>
</table>
DISCUSSION

The pathophysiology of the MTrP has become better understood as a result of recent studies on both human and animal subjects. Hong has hypothesized that there are multiple basic units in an MTrP region. He has speculated that each MTrP unit contains a sensitive locus (a minute site from which a LTR can be elicited when this site is mechanically stimulated) and an active locus (a minute site from which spontaneous electrical activity can be recorded). The sensitive locus is probably nociceptors (sensory component) and the active locus is possibly dysfunctional endplates (motor component). Spontaneous electrical activity (EPN) recorded from an active locus of MTrP has been theorized to be abnormal endplate potentials attributable to excessive release of acetylcholine, with the excessive acetylcholine release responsible for the formation of taut bands, which contain MTrPs. In an electron microscopic study, remarkable shortening of the sarcomeres in a contraction nodule (probably in or near the endplate zone) with lengthening of the sarcomeres outside the nodule was observed in some fibers (taut fibers). The total length of these muscle fibers was unchanged, but the tension of the muscle fibers was increased (taut band).

In previous studies, it has been suggested that the irritability of an MTrP is closely related to the LTR loci. The pressure pain threshold can be used for the assessment of MTrP irritability. In this study, we have demonstrated the high correlation between the irritability (measured with pain intensity and pressure pain threshold) and the prevalence of EPN loci in an MTrP region of upper trapezius muscle. It is our experience during EMG examination that when the EMG needle approaches a painful locus, EPN frequently can be recorded from this site. Therefore, an EPN locus may be in the immediate vicinity of an LTR locus. However, it is unclear whether every EPN locus in the MTrP region is closely connected with an LTR locus. Because the irritability of MTrP is related to the prevalence of both LTR loci and EPN loci, it is very likely that the vast majority of EPN loci are connected with LTR loci. Stated differently, in most endplates in the MTrP region, there are nociceptors (free nerve endings, as observed in animal study) in the nearby region. It has also been demonstrated that free nerve endings were frequently found near the sites from which EPN could be recorded. However, this does not indicate any neural connection between an EPN locus and an LTR locus.

It has been suggested that MTrP formation is related to a central mechanism on the basis of the studies on referred pain and LTRs. However, the mechanism of the connection between EPN loci and LTR loci is unknown. The correlation between them is possibly attributable to a peripheral interaction rather than a central sensory connection. It is possible that the persistent contraction of taut band may cause hypoxia and ischemia in the local region (within the endplate zone), which subsequently cause the release of inflammatory substance to sensitize the nociceptors (LTR loci). This is the peripheral sensitization of an MTrP. Gunn also suggested that local contraction of the muscle fibers (local depolarization without action potentials) may cause hypersensitivity of nociceptors. Therefore, the sensitivity of the nociceptors (irritability of MTrP) seems closely related to the prevalence of EPN, which could be related to the degree of local depolarization.

In conclusion, the irritability of an MTrP in the upper trapezius muscle is closely related to the prevalence of EPN in this MTrP region. In future research, assessment of EPN prevalence may be used for the estimation of MTrP irritability.
REFERENCES


Evaluation of Two Assessment Tools in Predicting Driving Ability of Senior Drivers

ABSTRACT

Objective: To evaluate Motor Free Visual Perceptual Test (MVPT) and Clock Drawing Task (clock test) as quick assessment tools in predicting driving capability of senior drivers for an on-road driving test.

Design: Senior drivers (≥55 yrs) referred for evaluation and recommendation for license renewal were given the MVPT, clock test, and an on-road driving test. Receiving operating characteristic (ROC) analysis and stepwise multivariate logistic regression (SMLR) were used to develop a probability model to differentiate between capable and incapable senior drivers.

Results: Data for 232 seniors who had completed all written tests and the on-road driving test were analyzed. Of the 232 seniors, 131 (56%) were classified as capable and 101 (44%) as incapable drivers on the road test. Mean scores for capable and incapable drivers were MVPT 32.0 ± 4.0 vs. 28.4 ± 4.6 and mean clock test score 3.5 ± 0.8 vs. 2.7 ± 1.2, and mean processing time was 7.1 + 6.5 vs. 10.6 + 5.5. The means of the three measurements were significantly different between the two groups (P value <0.001). ROC curve analysis revealed an optimal cut point of ≥32 for MVPT score with 60% sensitivity and 83% specificity. The optimal cut point for clock test scores is ≥3 with 70% sensitivity and 65% specificity. The optimal cut point for processing times is ≤6.27 secs with 60% sensitivity and 80% specificity. SMLR showed that the most significant predictor of seniors’ driving capabilities are the MVPT test scores and clock test scores.

Conclusion: MVPT and clock test tools are significant predictors of driving capability on an on-road driving test.

Key Words: Motor Free Visual-Perception Test (MVPT), Clock-Drawing Task (Clock Test), Processing Time, Cognitive Testing, Senior Drivers
Driving is an activity requiring complex sensory, perceptual, cognitive, behavioral, and motor skills. Crash rates increase steadily in senior drivers after the age of 60 yrs.\(^1\)\(^–\)\(^5\) Age-associated diseases and changes in physiologic function often combine synergistically to impair driving ability.\(^3\)\(^,\)\(^4\)

Senior drivers use automobiles for 90% of their outings.\(^6\) It is expected that there will be a threefold increase of senior drivers in the next few decades.\(^6\) By the year 2020, one of five aging baby boomers will be 65 yrs or older, and nearly half of these people will be more than 75 yrs old.\(^7\)

Senior drivers have a higher rate of accidents per mile than all other age groups, with the exception of the teenage group.\(^1\)\(^,\)\(^8\) The actual number of crashes senior drivers are involved in as a group is no higher than that of the general driving population because seniors drive fewer miles. When crash rates of senior drivers are adjusted for actual miles driven, their rate is higher than any age group, with the exception of drivers under the age of 25 yrs.\(^1\)\(^,\)\(^9\)–\(^18\) Senior driver crash rates per 100,000 miles driven are double those of the younger cohorts. Senior drivers also differ from younger age groups in the types of crashes they are involved in. Studies indicate that crashes involving senior drivers are more likely to occur at intersections and at low speeds, and they are five times more likely to involve two or more vehicles, with a higher frequency of fatality.\(^19\)–\(^22\)

Research is currently being conducted to make all aspects of driving safer and to determine how senior drivers with disabilities or unsafe drivers can be trained to drive safely. Some senior drivers who have deficiencies identified and corrected can become safe drivers.\(^1\) It is a challenge to safety officials to separate safe from unsafe drivers. The division of motor vehicles offices in most states are not adequately equipped to perform this task effectively.\(^1\)

### PHYSIOLOGIC CHANGES

Vision is very important for most of the sensory contributions during driving. With advancing years, multiple visual functions start to decline.\(^4\)\(^,\)\(^20\)\(^,\)\(^23\)\(^,\)\(^24\) Static visual acuity for licensing is a widely accepted standard. It is dynamic visual acuity (clarity of vision for objects moving relative to the observer) that shows a strong and consistent relationship to driving.\(^25\)\(^,\)\(^26\) There is a gradual progressive age-related decline in the nasal and temporal visual fields in the fifth decade.\(^25\)\(^–\)\(^27\)

Advancing age causes increased difficulty in night driving.\(^26\) Changes in age-associated lens protein and increased lens density cause increased scatter of light.\(^26\) The time required for glare recovery also increases with age.\(^20\)\(^,\)\(^26\) Increased lens density and decreased pupil size reduce illumination reaching retinal photoreceptors. Adaptation to darkness occurs more slowly in senior drivers. Higher levels of contrast are required to visualize objects adequately under conditions of low illumination.\(^20\)\(^,\)\(^26\)\(^,\)\(^29\)

Other factors that influence the visual system can result in several ocular diseases with additional reversible or progressive visual impairments.\(^26\) Some ocular diseases, including diabetic retinopathy, corneal diseases, cataracts, glaucoma, and age-related macular degeneration, can impede safe driving.

Hearing loss attributable to advancing age can cause high-frequency loss, which impairs the perception of important driving cues such as horns, sirens, and/or train whistles.\(^3\)\(^,\)\(^23\)\(^,\)\(^29\)\(^–\)\(^31\) Driving is also affected by age-related decline in musculoskeletal function (including decreases in strength, flexibility, coordination, or dexterity).\(^3\)\(^,\)\(^23\)\(^,\)\(^29\)\(^,\)\(^30\)

Age-related vision deterioration, decline in hearing and cognitive processing, and decreased limb strength and range of motion present special challenges to senior drivers, especially when driving at night, through intersections, on freeways, and in highway work zones.

Dynamic visual acuity (the ability to see moving objects) and scotopic visual acuity (the ability to see objects in dim light), contrast sensitivity (objects that do not stand out sharply from their background), or peripheral vision (objects that seem in the outer edges of visual field) are some components of good visual health essential for driving. In addition, attention, visual spatial skills, memory, information processing, and quick decision making are critical cognitive components for driving.\(^2\) Ongoing endeavors by the automobile industry to add more safety features to motor vehicles such as multiple air bags and navigational warning devices will benefit all age groups of drivers.\(^3\)\(^,\)\(^6\)

It would be beneficial for division of motor vehicles offices to use simple, inexpensive screening tools to differentiate capable from incapable drivers. Few states require repeated written and road tests for license renewal. Visual screening is limited to static visual acuity testing in almost all states.\(^5\)

### OBJECTIVES

The objective of this study was to evaluate the Motor Free Visual Perceptual Test (MVPT) and the Clock Drawing Task (clock test) as quick assessment tools in predicting driving capability of senior drivers for an on-road driving test.

### MATERIALS AND METHODS

#### Subjects/Participants

The subjects were 255 senior drivers, (\(\geq 55\) yrs of age) referred for a driver’s evaluation at a reha-
Clock Test

The clock test is designed to quickly assess visual–spatial construction, visual perception, and abstract conceptualization19,22,30–33 using three subtests: clock drawing, clock setting, and clock reading. The clock test measures the subject’s level of cognitive impairment and helps differentiate between normal elderly individuals and those suffering from dementia.29,30,32–34 There were conflicting results from another study that had limitations in detecting mild or early dementia, which was clinically detectable.35

The four-point method of clock drawing task by Pfizer, Inc.36 that is currently being used by the TriAD (Three for the Management of Alzheimer’s Disease: The Clinician, Patient, and Caregiver Alliance) was selected because it is easy to administer and easy to score. It requires only a paper and a pencil and is less ambiguous than conventional clock drawing tests.

The clock test provides well-defined, specific scoring criteria that ensure greater efficiency and reliability. Errors are broken down into the following components: (1) omissions, (2) distortions, (3) misplacements, (4) preservations, (5) substitutions, (6) additions, and (7) rotations.

The test subjects are instructed to draw the face of a clock and place the hands of the clock at numbers 10 and 11. The scores are given using the four-point system.36

The clock test was more sensitive than the clock drawing alone.32 It was based on a normative sample of elderly subjects aged 55 and older. Information from a large clinical sample was also used.29 The four-point scoring system for the clock test tool by Pfizer36 was used because it has high sensitivity and specificity (average 85%) and has a high level of interrater and test–retest reliability and predictive validity. It takes 5–10 mins to administer this test, including hand scoring for each component. The four-point scoring system was used for the clock test in this study.

MVPT

The MVPT assesses a subject’s visual perceptual ability, with no motor involvement needed to make a response. It is especially useful for individuals who have learning, motor, or cognitive disabilities.9–11,22 Five categories of visual perception are measured: spatial relationships, visual closure, visual discrimination, visual memory, and figure ground.

MVPT standards are based on a nationally representative sample that included senior drivers. Another feature of the MVPT is that the response time is standard, which is often useful in rehabilitation settings. The test format is visual multiple choice. There are 36 testing plates. The subject gets a point for each correct answer (Table 1). The processing time is measured by the response time for each plate viewed by the senior driver. The standard response time for the age group ≥50–69 yrs is 3–5.4 secs; for the age group ≥70–80 yrs, it is 4.5–7.1 secs for each plate viewed. The recorded response times in seconds are added and divided by 36, which gives the total processing time.

This task requires viewing all pictures/figures and looking for shapes and consistencies. The test provides a perceptual quotient and a perceptual age score. The MVPT takes approximately 25 mins to administer. In our study, average time to complete the MVPT was also evaluated as a predictor of driving capability. Scoring is extremely easy because raw scores can be quickly converted to standard scores and percentile ranks. Response-time data identify whether an individual’s responses are significantly delayed.12

The scores for this test range from 0 to 36, with a higher score indicating better visual perception. A maximum score of 36 indicates no errors. For the purposes of this study, raw scores were the primary outcome of interest because driving, regardless of age, requires specific visual perception skills.9 Both the clock test and MVPT have been used as screening tools to determine whether subjects are fit to undergo an on-road evaluation.9,29,32

In this study, MVPT was administered first, followed by the clock test. The on-road test was done on a different day in most of the senior drivers. The evaluator reviewed each driver’s MVPT and clock test scores before taking the driver for the on-road test.

bilitation center from September 1999 to October 2001 in northwest Ohio and southeast Michigan. Most of these subjects were from within the greater Toledo area (218/255). These subjects were referred for a driver’s evaluation at a rehabilitation center based on unsafe behaviors observed by law enforcement officers, including those that form a basis for citations; accidents; physician referral; referral by relatives; and self-referral. Senior drivers’ demographic information such as age, gender, place of residence, reason for referral, and details of citation (if that was the reason for referral) were also collected. These subjects were given two paper-and-pencil tests of cognitive functioning: the clock test and the MVPT.

The evaluator screened the subjects for any preexisting medical conditions that would legally preclude them from driving, such as visual homonymous hemianopsia (a primary visual impairment inadequately improved by corrective lenses), class IV cardiac status, and uncontrolled seizures.9

The clock test provides well-defined, specific scoring criteria that ensure greater efficiency and reliability. Errors are broken down into the following components: (1) omissions, (2) distortions, (3) misplacements, (4) preservations, (5) substitutions, (6) additions, and (7) rotations.

The test subjects are instructed to draw the face of a clock and place the hands of the clock at numbers 10 and 11. The scores are given using the four-point system.36

The clock test was more sensitive than the clock drawing alone.32 It was based on a normative sample of elderly subjects aged 55 and older. Information from a large clinical sample was also used.29 The four-point scoring system for the clock test tool by Pfizer36 was used because it has high sensitivity and specificity (average 85%) and has a high level of interrater and test–retest reliability and predictive validity. It takes 5–10 mins to administer this test, including hand scoring for each component. The four-point scoring system was used for the clock test in this study.

The MVPT assesses a subject’s visual perceptual ability, with no motor involvement needed to make a response. It is especially useful for individuals who have learning, motor, or cognitive disabilities.9–11,22 Five categories of visual perception are measured: spatial relationships, visual closure, visual discrimination, visual memory, and figure ground.

MVPT standards are based on a nationally representative sample that included senior drivers. Another feature of the MVPT is that the response time is standard, which is often useful in rehabilitation settings. The test format is visual multiple choice. There are 36 testing plates. The subject gets a point for each correct answer (Table 1). The processing time is measured by the response time for each plate viewed by the senior driver. The standard response time for the age group ≥50–69 yrs is 3–5.4 secs; for the age group ≥70–80 yrs, it is 4.5–7.1 secs for each plate viewed. The recorded response times in seconds are added and divided by 36, which gives the total processing time.

This task requires viewing all pictures/figures and looking for shapes and consistencies. The test provides a perceptual quotient and a perceptual age score. The MVPT takes approximately 25 mins to administer. In our study, average time to complete the MVPT was also evaluated as a predictor of driving capability. Scoring is extremely easy because raw scores can be quickly converted to standard scores and percentile ranks. Response-time data identify whether an individual’s responses are significantly delayed.12

The scores for this test range from 0 to 36, with a higher score indicating better visual perception. A maximum score of 36 indicates no errors. For the purposes of this study, raw scores were the primary outcome of interest because driving, regardless of age, requires specific visual perception skills.9 Both the clock test and MVPT have been used as screening tools to determine whether subjects are fit to undergo an on-road evaluation.9,29,32

In this study, MVPT was administered first, followed by the clock test. The on-road test was done on a different day in most of the senior drivers. The evaluator reviewed each driver’s MVPT and clock test scores before taking the driver for the on-road test.
On-Road Driving Test

The senior driver evaluation program is a copyrighted, standardized program of the Ohio State University Medical Center, Office of Geriatrics and Gerontology, administered through a license agreement with the Office of Geriatrics and Gerontology, the Ohio State University Health Sciences Center. The Lucas County Older Driver Evaluation Program was funded in part by the Ohio Department of Public Safety/NHTSA.

Three occupational therapists are trained as driver evaluators/trainers by the Ohio Department of Public Safety. Each occupational therapist is required to take a 40-hr course and additional 40 hrs of training to train/evaluate handicapped subjects. The MVPT was done first because subjects tend to get tired after some time and are unable to perform other tests. For consistency, only one evaluator conducted the on-road tests.

In the on-road test, the subject was allowed to get the feel of the car in the parking lot and then take it for a standard test drive. Each subject was expected to drive 17 miles in city traffic, highway traffic, and in residential areas, with 12 left turns, 5 right turns, 7 straight-through intersections, stop signs, and red lights at intersections with flashing red light or orange lights. They were observed for staying on course and within the lanes and lane changes, keeping within speed limits, and noting speed bumps. They were also taken to a busy parking lot, where each subject had to locate a parking spot and park with the car aligned correctly, then pull out of the parking lot. They were given three chances to correct themselves if they did not align the car correctly the first time.

Data Analysis

A commonly used and accepted level of certainty of 95% was set before the running of statistical comparisons. This standard preset level of certainty, or confidence, represents the likelihood that the results are true or correct and not attributable to random chance. For the result to be true, the alpha should be less than 0.05. The actual level of significance for any obtained statistic is a $P$ value $<0.05$.

$\chi^2$ analysis was used to evaluate the relationship between the categorical variables and forward exact test used for $2 \times 2$ categorical variables. One-way analysis of variance was used to compare the continuous variable means within the groups. ROC curve analysis, stepwise logistic regression analysis, and Hosmer–Lemeshow goodness-of-fit predicted probability tests were also performed. All statistical analyses were performed using Statistical Package for the Social Sciences, (SPSS) version 10.1 (SPSS Inc., Chicago IL).

Both the clock test and the MVPT were included in the model to identify their usefulness as predictors of the driving abilities of senior drivers (capable vs. incapable of driving) and to determine whether they could be used as quick indicators that a subject might need a more expensive, traditional driving test. The results of these subjects were discussed in a group meeting with a board-certified geriatrician and occupational therapists. The subjects’ test scores were used to categorize these subjects as capable or incapable drivers.

RESULTS

Two hundred fifty-five senior drivers were referred for evaluation and recommendation for license renewal. Referrals of these subjects were based on various violations or behaviors contribu-
ing to accidents. After excluding 23 subjects who did not complete all of the tests, data were analyzed for 232 senior drivers, who completed the MVPT, clock test, and on-road driving test.

There were 82 reported violations in 48 (59%) males and 34 (41%) females. In violations, females vs. males were no yield at stop: 2 4 vs. 17%; turning across traffic: 1 7 vs. 4%; running a red light: 9 vs. 13%; speeding: 3 vs. 15%. Other violations, including wrong way, yellow light, lane changes, slow speed, rear ending, and illegal U turn, were less frequent, with similar numbers in both genders (Table 2). No statistical analysis was done on violations (including gender).

ROC analysis and multivariate logistic regression were used to develop a predicted probability model to differentiate between senior drivers' capabilities. Subjects were categorized into two groups: 131 (56%) were capable drivers and 101 (44%) were found to be incapable drivers on the road test. The mean scores of the three measurements (MVPT score, clock test score, and processing time) in the capable vs. incapable drivers were significantly different between the two groups, with a $P$ value $<0.001$ (Fig. 1, Table 3 and 4).

ROC curve analysis revealed that the optimal cut point for the MVPT score was $\geq 32$ to yield 60% sensitivity with a confidence interval of 50.7–67.7 and 83% specificity and a confidence interval of 74.4–89.9, with a positive likelihood ratio of 2.95 and a negative likelihood ratio of 0.49. The optimal cut point for the clock test score was $\geq 3$ to yield 70% sensitivity with a confidence interval of 59.5–76.0 and 65% specificity and a confidence interval of 55.4–74.4, with a positive likelihood ratio of 1.97 and a negative likelihood ratio of 0.49. The optimal cut point for the processing time was $\leq 6.27$ secs to yield 61% sensitivity with a confidence interval of 52.6–69.5, and 79% specificity with a confidence interval of 70.0–86.6, with a positive likelihood ratio of 2.95 and a negative likelihood ratio of 0.49.

Stepwise multivariate logistic regression and the Hosmer-Lemeshow test of goodness fit were used to arrive at a predicted probability of driving capability. Figure 2 shows the relationship between predicted driving capability with MVPT score and clock test. MVPT score and clock test score were significant predictors of seniors' driving capabilities.

DISCUSSION

Driving is a controlled task involving cognition and operational, tactical, and strategic function. Senior drivers experience gradual declines in the sensory system, perceptual skills, and cognitive functioning with advancing age. Many seniors experience anxiety performing new tasks because of associated impairments, such as excessive risk-taking behavior, poor judgment, confusion, diminished frustration tolerance, and emotional instability.

Road tests evaluate practical driving skills and are a standard method of assessing driving competence. This study was designed to evaluate two off-road assessment tools that can be used to identify at-risk senior drivers before an on-road evaluation. The MVPT and the clock test were chosen to identify at-risk subjects. Senior drivers were referred to a rehab center for a driving diagnosis before recommending continuation of independent driving. The traditional way to diagnose driving capability is to administer a standard driving test. Several diagnostic tools have been sought as a replacement or an adjunct. However, earlier studies tested these tools individually. Although the results were encouraging, these tools cannot be used independently with reliable certainty.

This study was designed to evaluate the effectiveness of two simple assessment tools that...
can be easily administered in an office setting. The clock test and the MVPT were included in the model to identify whether they were useful predictors of driving diagnosis of senior drivers (capable vs. incapable of driving) and whether they could be used as quick assessment tools before recommending continuation of independent driving.

Traffic violations may vary by age group. Seniors, as a group, are highly cited in violations for

<table>
<thead>
<tr>
<th>TABLE 3 Analysis of MVPT score, clock score, and processing time in capable and incapable senior drivers</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tests</strong></td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>MVPT score</td>
</tr>
<tr>
<td>Incapable</td>
</tr>
<tr>
<td>Capable</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td>Processing time score</td>
</tr>
<tr>
<td>Incapable</td>
</tr>
<tr>
<td>Capable</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td>Clock test score</td>
</tr>
<tr>
<td>Incapable</td>
</tr>
<tr>
<td>Capable</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

MVPT, Motor Free Visual Perceptual Test.
failure to obey signs, failure to yield right of way, improper turns (particularly left turns), and illegal passing. Senior drivers have a tendency to make errors of omission (such as failure to read traffic signs or obey traffic lights) and demonstrate difficulty in maneuvers requiring a series of rapid judgments. A recent study showed that senior drivers fail to yield at stop signs and fail to yield to right of way. One study indicated that older female drivers have a higher incidence of traffic violations and have driving difficulty with increasing age.

Carr et al. analyzed the Fatality Analysis Reporting System from January 1990 to December 2000 to identify crashes based on age group (i.e., 25–64, 65–74, 75–84, and ≥85 yrs). Their study

| TABLE 4 Analysis of MVPT score, clock score, and processing time between groups and within groups of capable and incapable senior drivers |
|---------------------------------|--------|----------------|---|----------------|
|                                | Sum of Squares | df   | Mean Square | F     | P value |
| MVPT score                     |           |      |             |       |        |
| Between groups                 | 783.574   | 1    | 783.574     | 43.181| 0.000  |
| Within groups                  | 4300.665  | 231  | 18.146      |       |        |
| Total                          | 5084.240  | 232  |             |       |        |
| Processing time score          |           |      |             |       |        |
| Between groups                 | 676.208   | 1    | 676.208     | 18.313| 0.000  |
| Within groups                  | 8714.465  | 231  | 36.926      |       |        |
| Total                          | 9390.673  | 232  |             |       |        |
| Clock test score               |           |      |             |       |        |
| Between groups                 | 31.064    | 1    | 31.064      | 31.405| 0.000  |
| Within groups                  | 231.462   | 231  | 0.989       |       |        |
| Total                          | 262.525   | 232  |             |       |        |

Significant difference between two groups (capable vs. incapable) for the three scores (MVPT, processing time, and clock test).

FIGURE 2  Relationship between predicted driving capability and visual score (MVPT) and clock score.
showed that in-person license renewal in senior drivers significantly lowered fatality rates.\textsuperscript{43}

In this report, all subjects used the standard driving route, contrary to another report in which only 78\% of subjects used a standard driving route.\textsuperscript{44}

The current study analyzed MVPT, the clock test, and processing time scores. Fifty-six percent (131/232) of seniors were classified as capable drivers, with higher MVPT scores and clock test scores and lower processing time. Forty-four percent (101/232) of the seniors tested were found to be incapable drivers with lower MVPT scores and clock test scores and higher processing time on the standardized on-road driver evaluation tests. (Table 3). The means of the three measurements were significantly different between the two groups ($P < 0.001$; Table 4).

The mean MVPT score in this study was 32.0 ± 4.0 vs. 28.4 ± 4.6 (capable vs. incapable drivers), with 60\% sensitivity and 83\% specificity. The mean processing time was 7.1 ± 6.5 vs. 10.6 ± 5.5 secs, with 61\% sensitivity and 79\% specificity.

A recent study compared Cognitive Behavioral Driver’s Inventory (CBDI) scores for clients who passed and failed a driving evaluation. This study examined Cognitive Behavioral Driver’s Inventory, MVPT, Bells test, and on-road driving test results.\textsuperscript{38} The results of the on-road tests showed that 54\% of subjects passed their driving evaluations, 34\% failed, and 12\% were advised to take lessons and be reevaluated.\textsuperscript{38} The mean MVPT score in the pass group was 32.92; for the fail group, it was 29.89, with processing times of 4.63 and 6.11 in the fail group.\textsuperscript{38}

Earlier studies indicated a cutoff MVPT score of $\geq 30$ as the best predictor of a pass or fail on the on-road test; this cutoff is well accepted.\textsuperscript{9,45} The clock test score cutoff of $\geq 3$ is the best predictor;\textsuperscript{36} The clock test used in this study has a very simple four-point scoring system,\textsuperscript{36} whereas other tests have scoring systems with high score numbers. Because each assessment tool uses a slightly different scoring system, comparisons cannot be made.\textsuperscript{33}

The findings of earlier studies provide a guide for clinicians to identify the screening tool that can predict at-risk senior drivers. However, the findings indicate that different approaches may be required to accurately address the specific needs of each diagnostic group.\textsuperscript{38}

A working group of researchers from the United States and Canada are exploring the possibility of forming consensus guidelines on the basic components of off-road and on-road driving assessment.\textsuperscript{46,47} These guidelines will include key components that should be part of a standard assessment for determining the driving ability of senior drivers. The use of these guidelines will help stan-

dardize the assessment process regardless of which center subjects visit for their driving evaluation.\textsuperscript{44} Physicians should periodically look for any cognitive changes in their patients because these changes can impact their driving skills over time.\textsuperscript{48} A 6- to 12-mo repeat cognitive test should be planned by the physician. The AMA recommends two cognitive tests to identify at-risk senior drivers.\textsuperscript{59}

For senior drivers who perform poorly on off-road tests, a recommendation to an occupational therapy specialist is warranted for performance-based road testing.\textsuperscript{50} On-road evaluations by rehabilitative facilities are expensive ($\$200–$500) and may not be covered by medical insurance.\textsuperscript{51} For high-risk drivers, referral to the division of motor vehicles office can be considered for in-person license renewal.\textsuperscript{52}

**Implications of the Study**

This study provides easy-to-administer tools to healthcare providers to screen and identify at-risk drivers. Consensus guidelines, the MVPT, and the clock drawing tasks will be effective assessment tools to evaluate senior drivers for on-road tests.

**Limitations of the Study**

This was a retrospective analysis done on a variety of referrals. The study was not designed to prospectively analyze senior driver capability with and without underlying neurological conditions. Possible factors that may have reduced the association found between the MVPT, clock test scores, and on-road performance are limitations related to the evaluator’s knowledge of the test scores before administration of the on-road test. Gender differences were not analyzed in this study.

**CONCLUSION**

The current study indicates that the MVPT and the clock test tools were significant predictors of driving capability of senior drivers for an on-road driving test.

These tests have the potential to identify at-risk drivers. Physicians could use these simple, inexpensive tools (clock test and MVPT) in their clinical setting as screening tools before recommending seniors for independent driving.

**ACKNOWLEDGMENTS**

The authors thank Linda Mauger, Program Manager Office of Geriatric and Gerontology, Ohio State University; Shirley Mickens, Secretary, Total Rehab at Flower Hospital; Christine McKeen, Manager of Communications at The Toledo Hospital; and Yahya Daoud, Jobst Vascular Center for their help with the study and preparation of the manuscript.
REFERENCES

BOOK REVIEW

Total Knee Arthroplasty


The author has performed nearly 4000 primary knee replacements and has taught more than 500 orthopedic residents and arthroplasty fellows his surgical techniques. This book provides detailed instruction in the author’s total knee arthroplasty procedures. Surgical techniques are presented in the text and in the included DVD. The intended audience is primarily orthopedic residents or fellows, but the text would also be of interest to others who commonly participate in the care of patients who have undergone total knee arthroplasty.

The book contains 19 chapters. The first chapter reviews the history of total knee arthroplasty, with a focus on the procedures and materials developed in Boston, Massachusetts. Very well illustrated chapters follow that provide detailed instruction in the surgical procedures for a routine total knee arthroplasty as well as for complicated cases—varus or valgus deformity, knee stiffness, contracture, following osteotomy, for rheumatoid arthritis, and for bone stock deficiency. Unicompartmental knee arthroplasty is discussed as well as when bilateral procedures are indicated. Surgical complications are reviewed.

The book contains over 300 high-quality color surgical photographs, drawings, and radiographs illustrating the various steps of the surgical procedures. The narrated DVD demonstrates a total knee arthroplasty, a unicompartmental arthroplasty, and a lateral release for valgus.

Discussion of rehabilitative care is brief. The text does include a final chapter of frequently asked questions by patients about total knee arthroplasty. Useful patient-education material is found in this four-page chapter.

Overall, this is a well-illustrated and detailed surgical instruction guide for orthopedists in training. Physiatrists who work extensively with total knee arthroplasty patients would find it useful as a reference book to expand their knowledge of the surgical techniques used during these procedures.

Rating: *** for the average physiatrist; ***** for the orthopedist in training

John A. Schuchmann, MD
Temple, Texas
A Washing Toilet Seat with a CCD Camera Monitor to Stimulate Bowel Movement in Patients with Spinal Cord Injury

ABSTRACT

Objective: The objective was to study the effectiveness of a modified washing toilet seat equipped with a CCD camera monitor and an electronic bidet to facilitate precise hitting of the anal area with water streams to stimulate bowel movement in patients with spinal cord injury (SCI).

Design: There were 20 subjects, all of whom had traumatic SCI, were at least 5 mos post acute injury, and could change their position on the toilet seat while watching the monitor. Stimulation of bowel movement with the modified toilet seat was provided for a maximum of 30 mins. Success or failure to induce bowel movement was evaluated as related to injury level, ASIA impairment scale, and ability to voluntarily increase anorectal pressure, measured with a manometory.

Results: Bowel movement was successfully induced in 15 of the 20 patients (75%). Success was not related significantly to injury level, ASIA impairment scale, or ability to voluntarily squeeze. Compared with their usual manner of bowel management, for which they spent more than 30 mins, time needed for successful bowel movement was shortened in 11 of 13 patients. No complications were observed.

Conclusions: This preliminary study suggests that our newly modified washing toilet seat with a monitoring system could be a useful alternative for bowel management in patients with SCI.

Key Words: Bowel Dysfunction, Defecation, Anal Sphincter, Rectal Stimulation
Together with bladder dysfunction, bowel dysfunction poses a major problem in patients with spinal cord injury (SCI). Its successful management, although very important, is often very difficult to achieve. Problems related to defecation, which vary from fecal incontinence to chronic constipation, may have grave implications for the patients’ reintegration into society.1-3 In the Stockholm SCI study, 41% of participants rated bowel dysfunction as a moderate to severe life-limiting problem.4 In another study, 41% of patients with SCI spent more than 1 hr, and 12.5% spent more than 2 hrs per bowel-evacuation episode.5

There have been several studies reported in the literature describing countermeasures for these bowel problems. Bisacodyl suppository, enema, and stool softeners are commonly used as laxative agents.6 However, they are sometimes ineffective, and hemorrhoids or rectal abscesses can occur among those who use primarily suppositories and enemas routinely. Christensen et al.7 have described the enema continence catheter (ECC) and Malone antegrade continence enema (MACE). The ECC is a catheter with an inflatable balloon; the catheter is inserted into the rectum, and the balloon is inflated to hold it there. The rate of success with the ECC is reported to be 57%. The MACE is a simple operative procedure that brings out the appendix to skin level to form an appendicostomy. Through the stoma, patients can introduce a catheter and administer an enema. Overall success with the MACE was found in 87% of patients.

Puet et al.8,9 have described pulsed irrigation evacuation, which uses intermittent, rapid pulses of warm water to break up stool impactions and stimulate peristalsis. The pulsed irrigation evacuation has been reported to be a safe and effective method of treating impactions. These devices, however, require assistance for administration, and the ECC and MACE entail the potential risk of bowel perforation when inserting a catheter into the anorectum or stoma.

To avoid these complications and to facilitate self-management of bowel care by the patient, we developed a less invasive procedure for rectal stimulation that uses a washing toilet seat equipped with a CCD camera monitor. The purpose of this study is to test the effectiveness of our modified bowel-care system to counter these difficult bowel problems in patients with SCI.

METHODS

We recruited 20 patients with traumatic SCI who received inpatient care at the SCI units of Murayama Medical Center in Tokyo, Japan, during the period from November 2002 to September 2003. The study design was a single test per subject. All of them were males, and their mean age was 46.3 yrs (SD = 17.9, range = 18–73). The level of injury was cervical in 11 patients, thoracic in 7 patients, and lumbar in 2 patients. The impairment level, as classified according to the American Spinal Injury Association (ASIA),10 was A in eight patients, B in four patients, C in four patients, and D in four patients. The patients were at least 5 mos post acute injury and could independently transfer to the toilet seat and change their position on it while watching the CCD monitor. The study was approved by the institutional ethics review board of our facility. The purpose and procedures of the study were fully explained, and informed consent was obtained from the participants.

The newly developed procedure to induce bowel movement uses a washing toilet seat equipped with an electronic bidet (Hygiene toilet seat, DL-GT30, Panasonic, Japan), a CCD camera (KPC-S20CB, KT&C Corp., Korea), and a light (DOP-151 Asahi electric Corp., Japan) (Fig. 1). The electronic bidet provides water flow. The water stream is pulsatile and adjustable and can deliver 0.55 liters of water and 0.6 liters of air per minute maximally. To facilitate precise location of the anorectal area, we devised a washing toilet seat equipped with a CCD camera monitor. By watching the monitor, the patients could accurately direct the water stream toward the anal area. The maximum duration of stimulation was set at 30 mins, and the maximum power of the water stream that the toilet seat could provide was used for stimulation. After 30 mins, the amount of residual stool in the rectum was examined using digital evacuation. During the study period, the participants kept taking stimulant laxatives (senna, magnesium oxide, alosenn, sennoside) if they were prescribed, but suppositories and enemas were discontinued.

To evaluate anorectal functions, anorectal manometry was performed in 18 patients. In left lateral decubitus position, a balloon catheter was inserted 15 cm deep from the anal canal, and it was pulled out slowly. First, we measured anal resting pressure, which represented the internal anal sphincter pressure, and then we assessed squeeze pressure at the external anal sphincter level during voluntary contraction. We expected that the water stream was likely to have invaded into the anal canal if voluntary contraction was difficult and if sphincter tone was reduced, which was likely to have caused reflex defecation. Then, we studied the relationship between the success of bowel evacuation and the ability to voluntarily squeeze. We related success or failure to induce bowel evacuation with injury level, ASIA impairment scale with $\chi^2$ test, and the ability to voluntarily increase anorectal pressure measured with anorectal manometry with a forward exact probability test. All analyses were performed with Statcel (OMS, Saitama), a
statistical software program for Windows, and the level of significance was set at $P < 0.05$.

RESULTS

Figure 2 shows the usual manner of bowel care practiced by our study participants. Sixteen patients (80%) used rectal suppositories, three (15%) used enemas, and one (5%) used Valsalva maneuvers. Fifteen of the 20 patients (75%) also used manual evacuation in addition to a routine bowel program. As illustrated in Figure 3, seven patients (35%) spent less than 30 mins for usual defecation, 12 (60%) spent 30–60 mins, and one (5%) spent more than 60 mins. Fourteen of the 20 patients (70%) could not direct the water stream precisely to the anorectal area without the use of the monitor. All patients in the ASIA impairment levels A and B ($n = 12$) and two of the four level C patients belonged to this category. On the other hand, two of the four level C patients and all of the level D patients could direct the water stream to the anorectal area without the monitor.

With their usual manner of bowel care, bowel movement was induced within 30 mins in seven patients, from 30–60 mins in 12 patients, and in more than 60 mins in one patient. In contrast, bowel movement was successfully induced with our modified device within 30 mins [average time $= 17$ mins (range 3–28)] in 15 of the 20 patients (75%). Time needed for bowel management with our modified device was shorter than that with patients’ usual manner of bowel care ($\chi^2$ for independence test, $P = 0.01$). Time required for successful bowel movement was shortened in 11 of 13 patients (85%) who had spent more than 30 mins with their usual manner of bowel management. However, residual stools were found in 8 of the 15 patients.

The success of defecation was not related significantly with injury level ($P = 0.69$) or ASIA impairment scale ($P = 0.44$) ($\chi^2$ for independence test).

The manometric results showed that the average internal anal sphincter pressure was 25.4 mm H$_2$O (SD 14.0, range 4–51), and the average external anal sphincter pressure was 43.3 mm H$_2$O (SD

FIGURE 1 System of a washing toilet seat with a CCD camera monitor. By watching the monitor, the patients could accurately direct the water stream toward the anal area. A, whole image; B, close-up of CCD camera and light.

FIGURE 2 Usual bowel-care methods used by the participants ($n = 20$).
180) mm Hg in healthy controls. Although we did not compare pressure in patients with SCI with that of healthy controls, our data indicate that the resting anal pressure and squeeze anal canal pressure were lower than normal values previously reported.

We expected that water stream was likely to have invaded into the anal canal if voluntary contraction was difficult and sphincter tone was reduced, which was likely to have caused reflex defecation. However, the success rate of defecation in patients who could not voluntarily squeeze (10 of the 12 patients, 83.3%) did not significantly differ from that in patients who could not voluntarily squeeze (10 of 12 patients, 83.3%). This might indicate that bowel movement in patients with SCI does not depend on anal sphincter pressure alone. However, the statistical nonsignificance was more likely to have been caused by the small sample size in the present study, and a larger study is needed.

Regarding complications, the risk of bowel perforation must be considered when using ECC or MACE. For manual evacuation, hemorrhoids associated with defecation and anorectal injury are major complications. In contrast, our device only applies water to the anal area, and it is unnecessary to insert anything such as a rectal suppository, enema, or catheter into the rectum. This may be one important reason why we experienced no complications in this preliminary study.

As for the time needed to induce bowel movement, 41% of the patients spent more than 1 hr, and 12.5% spent more than 2 hrs for bowel evacuation with their usual method of bowel care. With our new method, it was shortened to less than 30 mins. However, we must note that in 8 of the 15 patients who succeeded in inducing bowel movement within 30 min, residual stools were found. This indicates that more than 30 mins may be needed to completely empty the rectum. This is an area for future research.

In conclusion, although this study was a pilot study, and residual bowel stools were found in eight patients, our newly devised bowel-management system, which uses a washing toilet seat with a CCD camera monitor, successfully induced bowel movement in 75% of the patients with SCI in less time than was needed for their routine manner of bowel care. With our new method, it was shortened to less than 30 mins. However, we must note that in 8 of the 15 patients who succeeded in inducing bowel movement within 30 min, residual stools were found. This indicates that more than 30 mins may be needed to completely empty the rectum. This is an area for future research.

REFERENCES

**BOOK REVIEW**

**Anatomical Guide for the Electromyographer: The Limbs and Trunk**


This is the fourth edition of a series begun by Edward F. Delagi, MD, Aldo Perotto, MD, John Iazzetti, MD, and Daniel Morrison, MD.

Obviously, in a book such as this one, the main subject matter does not change, and there is a lot of repetition from previous editions. Nevertheless, the book is thoroughly updated and does include some new information that is very valuable. This book, as most electromyographers already know, contains descriptions of needle localization for many different muscles. For most muscles that are routinely tested in electromyography, the practitioner obviously does not need to look things up in a reference book, but occasionally a muscle needs to be tested that is not routinely tested, so a book such as this one is extremely valuable in looking at the techniques that are available.

I was very impressed with this book. It is a great reference manual. There are numerous small sections such as hand, forearm, arm, shoulder joint, etc. This makes looking up muscles easy. For each muscle, there is a useful description of the innervation path. What is particularly useful is that it does not just give the end nerve that innervates the muscle, but the entire course, through the plexi as indicated. The book’s inclusion of the origin and insertion of each muscle is also useful. There is additional information on pitfalls, and there are comments on how to test each muscle, as well as the electrode-insertion technique. For each muscle, there is also a cross-sectional view that shows where the muscle stands in relation to the other anatomic structures. The quality of those cross-sectional figures is not as nice as the other anatomic drawings that show the muscle course.

Compared with other books of its type, this guide contains more information; this also makes it a bit larger and less portable. Some of the extra information is not always needed. On balance, though, it is a solid work.

**Rating:** *****

Ralph M. Buschbacher, MD
Indianapolis, Indiana
Body Image in People with Lower-Limb Amputation
A Rasch Analysis of the Amputee Body Image Scale

ABSTRACT

Objective: The aim of this study was to examine the psychometric properties of the Amputee Body Image Scale (ABIS) through Rasch analysis, investigating the quality of its rating categories and its reliability and validity.

Design: The ABIS (20 items; ratings of 1–5) and Trinity Amputation Prosthesis Experience Scales (TAPES) were administered by post and completed by 145 people with a lower-limb amputation and currently wearing a prosthesis.

Results: According to Rasch analysis and expert review, some response categories were collapsed and six items were deleted. The remaining 14 items created a revised ABIS (ABIS-R) rated with a three-level rating scale. ABIS-R fitted the unidimensional construct that the scale was intended to measure and demonstrated good reliability (Cronbach’s alpha and person separation reliability = 0.87), targeting, and internal construct validity. Moreover, the correlations with the nine TAPES subscales (in particular, $r = -0.54$ with the general adjustment, $r = -0.43$ with the social activity restriction, and $r = -0.40$ with social adjustment) supported the convergent validity of ABIS-R.

Conclusions: The 14-item ABIS-R demonstrates good psychometric characteristics for measuring body image disturbances in people with lower-limb amputation. These preliminary results suggest the general adequacy of the new instrument and provide a good foundation on which further validation and psychometric studies of the ABIS-R can be conducted.

Key Words: Amputation, Body Image, Psychometrics, Rasch Analysis
Psychosocial factors have recently been demonstrated to influence the prosthetic rehabilitation of individuals with an amputation. Consequently, the psychological impact of amputation and the subsequent fitting of a prosthesis needs to be taken into consideration in conjunction with the physical impact to enhance the prosthetic experience and outcomes for the user and to lead to enhanced health and quality of life. For example, there are a number of images for the person who has experienced an amputation to adjust to: the “complete” or familiar body before the limb loss, the traumatized body, the healing body, and the extended body (i.e., a body supplemented with prosthetic devices and, if necessary, mobility aids). Each of these images may be accompanied by phantom sensations and/or phantom limb pain (i.e., sensations and/or pain in the part of the body that has been amputated). Rybarczyk et al. proffer that the person has to adapt to an image of him- or herself without the amputated limb while reconciling three images of his or her body: before the limb loss, without a prosthesis, and with a prosthesis.

Body image anxiety has been found to be significantly related to depression, poorer perceived quality of life, lower levels of self-esteem, higher levels of general anxiety, lower levels of prosthesis satisfaction, and lower levels of participation in physical activity. According to Horgan and MacLachlan, adaptation to a changed body image is a potential measure of psychosocial adjustment to amputation. Therefore, it is an important construct that should be included in a clinical assessment.

However, to be adequately able to investigate body image concerns in people with an amputation and to explore its relationship with psychosocial adjustment, it is important to have a psychometrically sound instrument that will facilitate the development of a solid evidence base on which to set up appropriate interventions. The Amputation Body Image Scale (ABIS) developed by Breakey has been proposed as one such instrument. It comprises 20 items that assess how a person with an amputation perceives and feels about his or her body experiences. In the original paper outlining its development, acceptable content validity, internal consistency, and convergent validity were demonstrated using classical test theory, and significant positive correlations were found between the perception of body image (using the ABIS) and psychosocial well-being (more specifically, anxiety, depression, self-esteem, and life satisfaction). In addition, Murray and Fox observed moderate to high negative correlations between body image disturbance as measured by the ABIS and prosthesis satisfaction. The validity of the ABIS was also supported by Wetterhahn et al., who found a significant correlation between six subscales of the Multidimensional Body–Self Relations Questionnaire and the ABIS. To the best of our knowledge, these are the only published studies using the ABIS. As such, many issues related to the structure and main psychometric properties of the scale still need to be examined. In particular, it was considered worthwhile to further investigate (a) content validity (i.e., that it covers all parts of the universe of content and reflects the relative importance of each part, and that it is free from the influence of factors that are irrelevant to the purpose of the measurement), and (b) the rating-scale structure.

Recently, there has been a growing trend in the field of rehabilitation to implement Rasch analysis to facilitate the development and validation of questionnaires. Rasch analysis provides psychometric information that is not given by classical test theory, examining, among other things, (a) how the rating scale is being used; (b) the validity of a measure by evaluating the fit of individual items to the latent trait; it postulates that if the ability in responding to items on an ordinal scale is explained by an underlying unidimensional construct, the hierarchy of difficulty of the items is expected to match the hierarchy of ability of the subjects (i.e., more able subjects are more likely to pass more difficult items) within a probabilistic model; and (c) whether the pattern of item difficulties is consistent with the expectations of the construct and, hence, provides an adequate description of the range and hierarchical relationship of the variable. Indeed, Andresen recommends Rasch analysis as a method for assessing scaling properties in addition to traditional psychometric criteria for reviewing and assessing surveys and questionnaires for disability outcomes research.

To facilitate the availability of a psychometrically robust measure of body image in people with an amputation, the aim of this study was to perform an in-depth validation of the basic measurement properties of ABIS through Rasch analysis, investigating the quality of their rating categories and the validity (unidimensionality and internal construct validity) and reliability of the instrument.

### METHODS

#### Procedure and Sample

After ethical approval from two national limb-fitting centers in Ireland, hospital charts of potential participants were reviewed. A preselection criterion included a requirement that the participants with loss/absence of a limb be at least 18 yrs old. A cover letter, the questionnaire, and a stamped, addressed envelope were sent to all participants. A
short reminder card was sent 2 wks after the initial mailing to nonrespondents. There was no incentive or reimbursement for participation.

One hundred ninety-one people returned completed questionnaires. Of these 191 respondents, 145 indicated that they had had a lower-limb amputation and were currently using a prosthesis. These respondents were included in the study. The characteristics of the sample (n = 145) are outlined in Table 1. As can be seen, the sample was predominantly male, with the prevalent cause of amputation being peripheral vascular disorder, diabetes/peripheral vascular disorder, or accident/trauma. In addition, the most common level of amputation was below the knee. These characteristics are consistent with the general population of people with lower-limb amputations in the Western world.

Measures

ABIS

The ABIS12 comprises 20 items (see Appendix) that assess how an amputee perceives and feels about his or her body experiences. Participants are asked to indicate their responses to the questions using a rating scale of 1 (none of the time) to 5 (all of the time). Three questions (3, 12, and 16) are reverse scored. The scale produces scores that range from 20 to 100, with high scores indicating high body image disturbance.

Trinity Amputation and Prosthesis Experience Scales

As well as requesting demographic and disability-related data regarding gender, age, cause and type of amputation, length of time living with the prosthesis, and degree of prosthetic use, the Trinity Amputation and Prosthesis Experience Scales (TAPES) consist of nine subscales.7 There are three psychosocial subscales: general adjustment (e.g., I have adjusted to having an artificial limb), social adjustment (e.g., I don’t mind people asking about my artificial limb), and adjustment to limitation (e.g., Being an amputee means that I can’t do what I want to do). Each of these subscales contains five items, which are measured along a five-point rating scale (strongly disagree, disagree, neither agree nor disagree, agree, strongly agree). Scores range from 5 to 25, with higher scores indicating greater levels of adjustment. The TAPES also contains three activity-restriction subscales: functional activity restriction (e.g., walking 100 yards), social activity restriction (e.g., visiting friends), and athletic activity restriction (e.g., sport and recreation). Each of these activity-restriction subscales contains four items, which are measured along a three-point scale (not at all limited, limited a little, limited a lot). Scores range from 3 to 12, with higher scores being indicative of greater activity restriction. There are three additional subscales assessing satisfaction with the prosthesis, measured along a five-point scale (very dissatisfied, dissatisfied, neither dissatisfied nor satisfied, satisfied, very satisfied). The functional satisfaction subscale contains five items (e.g., reliability), with a potential score range of 5–25. There are four items in the aesthetic satisfaction subscale (e.g., color), with a potential score range of 4–20. Because weight satisfaction contains only one item, scores in this subscale range from 1 to 5. Higher scores in each of the satisfaction subscales are indicative of greater satisfaction with the prosthesis. Each of the psychosocial, activity-restriction, and satisfaction scales demonstrate high internal reliability using Cronbach’s alpha (range, 0.75–0.89) and good face, content, construct, and predictive validity.7

The TAPES also looks at the experience of phantom limb pain, residual limb pain, and other medical problems not related to the amputation. Each of the aforementioned is subdivided into questions relating to (1) whether that type of pain is experienced (2), how often it is experienced, (3) how long each episode lasts, (4) how the level of pain can be described, and (5) the extent to which it interferes with their daily lives. This section of the TAPES also incorporates two items requesting respondents to rate their general health and phys-
ical capabilities, using a five-point scale (very poor [1] to very good [5]).

**Statistical Analysis**

A three-stage process was used to investigate the basic psychometric properties of the ABIS:

1) Using SPSS version 11 (SPSS 11 for Windows, SPSS Inc., Chicago, IL), internal consistency and homogeneity of the original 20 items were examined by calculating:

   a) Cronbach’s coefficient alpha. It has been suggested that this should be above 0.70 but not higher than 0.90,\(^{26-27}\) because it is important to strike a balance between satisfactory internal consistency and an instrument that is too homogenous and thus measures a very restricted aspect of a phenomenon.\(^{17,28}\)

   b) Item–total correlations. Spearman rank correlation coefficients (\(r_s\)) were used to examine the degree to which each item was correlated with the total score, omitting that item from the total. The usual rule of thumb is that an item should correlate with the total score with a correlation coefficient larger than 0.5.\(^{26-27}\)

   c) Kaiser Meyer Olkin measure of sampling adequacy (both global and for each item). The level of homogeneity of the matrix of item scores was investigated through comparison of the magnitudes of the correlation coefficients observed with those of the partial correlation coefficients. Kaiser Meyer Olkin values greater than 0.6 are sought because values less than this indicate that one or more items should not be included in the factor analysis, because they do not belong to the same universe shared by the other variables.\(^ {29}\)

   d) Principal-component analysis. To test the unidimensionality of the measure, the amount of variance explained by the first principal component and the extent to which the first eigenvalue was larger than the second and third ones is examined. It has been suggested that when the first component accounts for 40% of the total variance, it can be said that a set of items is measuring a single dimension.\(^{30,31}\)

2) Rasch analysis: The matrix of single raw scores for each subject was subjected to the rating-scale model through the WINSTEPS software (WINSTEPS, Chicago, IL) to estimate the following issues:

   a) Rating-scale diagnostics. To investigate whether the rating scales of the questionnaire were being used in the expected manner, the following criteria (suggested by Linacre\(^ {32}\)) have been adopted to judge this parameter: (1) at least ten cases per category; (2) regular distribution of category use; and (3) monotonic increase in both average measures across rating-scale categories and thresholds. Thresholds (sometimes also called step calibrations) are the points at which the probability of a response of either the category below and the next category are equally likely; that is, they represent the transition from the category below to the next category; (4) category outfit mean square values less than 2; and (5) thresholds differences greater than 1.4 and less than 5.

   Until all these criteria were met, categories were combined according to specific guidelines, and several categorizations have been compared, looking not only at above indicators of category diagnostics but also at various reliability and validity indices.\(^ {24}\) Only the final solution is reported on in the results.

   b) Validity. After the check and revision of the rating scale, validity was analyzed by evaluating the fit of the individual items to the latent trait (unidimensionality), as well as by examining whether the pattern of item difficulties was consistent with the expectation of the construct. Depending on the string of responses provided by a particular sample of subjects on a particular sample of items, the Rasch model estimates (1) goodness of fit of the model (fit statistics). Information-weighted (infit) and outlier-sensitive (outfit) mean-square statistics (MnSq) for each item were calculated (similarly to a \(\chi^2\) analysis) to test whether there were items that did not fit with the model expectancies. According to the literature, items are deemed to have an acceptable fit to the model when the MnSq is greater than 0.6 and less than 1.4.\(^ {24,33}\) Items outside this range were considered misfitting (MnSq > 1.4) or overfitting (MnSq < 0.6) (see also below); (2) the level of difficulty achieved by each item on an interval scale (item difficulty) and where each individual subject fits along the continuum (subject ability). Item difficulty and patient ability are expressed (on a common interval scale) in logits, the natural logarithm of the odds of mutually exclusive alternatives (e.g., pass vs. fail, or higher response vs. lower response).\(^ {24,34,35}\) It is reported that a sample size of about 100 persons will estimate item difficulty with an alpha of 0.05 to within 0.5 logits.\(^ {36}\)

   c) Reliability. The final set of items satisfying both the model fit requirements and an expert review were evaluated in terms of “separation” (G), defined as the ratio of the true spread of the measures with their measurement error.\(^ {24}\) The item separation index indicates an estimate (in standard error units) of the spread or separation of items along the measurement construct; the person separation index provides the same for persons (describing the number of performance levels the test measures in a particular sample). A separation of 2.0 is considered good and enables the distinction of three groups or strata (high, average, low) (number of distinct strata = \((4G + 1)/3\)) , defined as segments whose centers are separated by distances...
greater than can be accounted for by measurement error alone. A related index is the reliability of these separation indices, indicating the degree of confidence we can place in the consistency of our estimates (range = 0–1; coefficients >0.80 are considered good, and coefficients >0.90 are excellent).  

3) Construct validity: Both the original 20-item ABIS and the new Rasch-refined version of the scale (known as ABIS-R) were correlated with the TAPES using Spearman's correlation coefficient to assess their convergence issues. It is expected that the ABIS and ABIS-R will correlate positively with each of the activity-restriction subscales on the TAPES—that is, body image disturbance will be associated with higher levels of restriction in athletic, social, and functional activity restriction. It is expected that the ABIS and ABIS-R will correlate negatively with each of the satisfaction and adjustment subscales of the TAPES—that is, body image disturbance will be associated with lower levels of functional, weight, and aesthetic satisfaction and lower levels of general, social, and limitation adjustment. Aside from the significance level of the correlations, $r < 0.25$ indicates little relationship, and $r$ from 0.25 to 0.50 suggests a fair degree of relationship.

**RESULTS**

**Internal Consistency/Homogeneity**

The Cronbach's coefficient alpha of the original ABIS was 0.90. The item-to-total correlation coefficients ($r_{it}$) ranged from 0.30 (item 7) to 0.74 (item 11). The overall Kaiser Meyer Olkin measure of sampling adequacy was 0.87. The Kaiser Meyer Olkin statistics for each individual item found on the diagonal of the antiimage correlation matrix were all greater than 0.6, as required; they ranged from 0.94 (item 8) to 0.73 (item 15). Principal-component analysis showed that four factors had eigenvalues greater than 1.0, and the first factor (eigenvalue = 7.40) explained 37% of variance, whereas the second (eigenvalue = 1.61) and third (eigenvalue = 1.28) factors explained only small additional amounts of variance (9 and 6%, respectively). The scree test identified only the first factor before the “break.” The internal consistency indices and the large amount of variance accounted for by the first principal factor indicated that the test was sufficiently unidimensional to be submitted to Rasch analysis.

**Rasch Analysis**

Regarding rating-scale diagnostics, the rating categories of 2 (rarely), 3 (some of the time), and 4 (most of the time) did not comply with the set criteria for category functioning (average measures, thresholds, etc.) (Fig. 1A).

The model meeting all the established criteria and with the best person separation and reliability was the one that collapsed into a unique category level 2 with 3, and level 4 with 5 of the ABIS, thus producing a new three-level rating scale (0 = none of the time, 1 = sometimes, and 2 = most/all of the time). Figure 1B shows category probability curves after collapsing the categories.

After the phase of rating-scale modification, the Rasch analysis showed that 15 ABIS items fitted the unidimensional construct that the scale was intended to measure (Table 2). Items 2, 19, 17, and 16 were misfitting (MnSq > 1.4), and item 11 was overfitting (MnSq < 0.6).

Regarding the hierarchic ordering of items, Figure 2A shows the distribution map of subject ability and item difficulty of the items, according to the Rasch model. In the original 20-item scale (Fig. 2A), item measures ranged from $-1.63$ (item 19) to $+1.33$ logits (item 5), and person-ability measures ranged from $-3.00$ to $+4.19$ logits. In the shortened 14-item version without misfitting and overfitting items or item 7 (which was deleted for content reasons; see Discussion) (ABIS-R), item measures ranged from $-1.47$ (item 12) to $+1.23$ logits (item 5) (Fig. 2B), and person-ability measures ranged from $-3.21$ to $+3.99$ logits.

In the 14-item ABIS-R, the item separation index was 4.59 (item separation reliability = 0.95), the person separation index was 2.33, and the person separation reliability (which is analogous to Cronbach’s alpha) was 0.84.

**Correlations of ABIS (20 Items) and ABIS-R (14 Items) with TAPES**

As can be seen in Table 3, there is a significant negative correlation between both ABIS and ABIS-R and each of the following TAPES subscales: aesthetic, weight, and functional satisfaction; and general, social, and limitation adjustment. There is also a significant positive correlation between the ABIS, ABIS-R, and functional and social activity restriction.

**DISCUSSION**

Being cognizant of the person's body image after amputation is an important aspect of ongoing postamputation care, and its inclusion as part of a comprehensive outcome measurement merits consideration. However, because such measurements need to demonstrate robust psychometric properties, this study used Rasch analysis to improve the value of the ABIS by refining its rating scale, distinguishing items belonging to the same construct (unidimensionality), verifying the expected difficulty hierarchy of its items, examining the extent to which the items are of appropriate difficulty for the model meeting all the established criteria and with the best person separation and reliability was the one that collapsed into a unique category level 2 with 3, and level 4 with 5 of the ABIS, thus producing a new three-level rating scale (0 = none of the time, 1 = sometimes, and 2 = most/all of the time). Figure 1B shows category probability curves after collapsing the categories.
the sample (targeting), and analyzing reliability in terms of both internal consistency and separation.

**Internal Consistency**

The internal consistency of ABIS was found to be satisfactory according to multiple tests from classical test theory.\(^{27,28}\) Furthermore, despite the reduction of six items, the ABIS-R showed an acceptable value (0.84) of person separation reliability, which is analogous to Cronbach’s alpha.\(^{24}\)

**Rasch Analysis**

Regarding the rating-scale diagnostics, the ABIS showed some disordered thresholds. This contradicts the usual interpretation of categories (i.e., that they represent the sequence of the most

---

**FIGURE 1** Category probability curves of ABIS. The y-axis represents the probability (0–1) of responding to one of the rating categories, and the x-axis represents the different performance values (person measure minus the item measure) in logits. The 0 curve declines as the subject’s ability increases; the crossing point (where 0 and 1 are equally probable) is the first threshold. The same applies for the other curves. The plot should look (as in panel B) like a range of hills. A, Original scale with five categories (1 = none of the time, 2 = rarely, 3 = some of the time, 4 = most of the time, and 5 = all of the time). B, Revised scale after collapsing some categories (1 → 0 = none of the time; 2 + 3 → 1 = sometimes; and 4 + 5 → 2 = most/all of the time).

Apparent from Figure 1A is that the probability of using the central categories 2–4 is never higher than receiving other ratings. Conversely, in Figure 1B, it can be seen that the probability of selecting one of the three revised rating categories is now a clear function of the level of body image disturbance expressed by the individual in x-axis.
likely outcomes) and suggests that respondents were not able to distinguish their abilities as finely as envisaged by the five response categories (i.e., more categories existed in the scale than were needed to describe the construct). In particular, category diagnostics evidenced the inability of respondents to appreciably discern between categories 2 (rarely) and 3 (some of the time) and between categories 4 (most of the time) and 5 (all of the time) and later confirmed the appropriateness of collapsing them into two single categories indicating “sometimes” and “most/all of the time.” This modification makes sense, eliminates the redundancy of underused rating categories, and ensures that each rating category is distinct from the others. Similar problems and solutions with scales using adjectival descriptions of frequency and discrete responses have been illustrated by Bond and Fox (p 166) and Zhu.

Four ABIS items proved to be “misfitting”: their values suggested erratic response patterns. The misfitting could have been attributable—among other reasons—to their being part of another construct, being poorly written, or being too sensitive to confounding factors. We used a careful approach to item deletion that was based not only on the criterion of Rasch fit statistics but also on expert analysis of the item content and presentation, to guarantee the face, content, and clinical validity of the shortened scale. That stated, we suggest the elimination of the four misfitting items (because of their major threat to validity) for both statistical and content reasons. They are:

- Item 2: “I avoid wearing shorts in public because my prosthesis would be seen”; and item 17: “I wear baggy clothing in an attempt to hide my prosthesis.” The wearing of clothes may be influenced by factors other than body image. This task is also sensitive to cultural and environmental (including geographical) factors.
- Item 16: “I like the appearance of my stump anatomy.” It is possible that the way in which this item is phrased may render it sensitive to confounding factors; for instance, a significant proportion of people experience stump pain. However, the shape and appearance of the residual limb is a potentially important aspect of body image, and this element is retained in item 20, which remains in the ABIS-R.
- Item 19: “It is important that my prosthesis and remaining anatomy of my affected limb are the same size as the other limb.” We do not think that the element of “body size distortion” is as relevant to people with an amputation, nor is it a significant part of their concept of body image.

One item was “overfitting” (MnSq < 0.60)—that is, open to interpretation as redundant or failing to discriminate persons with different levels of ability. Overfitting items contribute little extra information beyond that provided by other items in the scale. However, the decision to remove, retain, or substitute overfitting items should be made on the basis of clinical reasoning. In our opinion, because the aspects covered in the overfitting item (item 11: “The loss of my limb makes me think of myself as disabled”) are also covered by remaining items, we would suggest removing it from the scale. Indeed, concerns about the reduction of functional capabilities are also found in items 4, 9, 13, and 15 (see Appendix). Furthermore, we suggest the removal of item 7 (“I experience a phantom limb”) on conceptual grounds. Because the majority of people with an amputation experience phantom limb pain, this item is unlikely to distinguish between people who are and are not experiencing body image disturbance. Moreover, this item was that with the lowest item–total correlation (r = 0.30), and some authors consider as satisfactory only a correlation of 0.40 or higher.

| TABLE 2 Item calibrations (measure) with standard errors (SE) and infit and outfit mean-square statistics (MnSq) for the 20 items of Amputee Body Image Scale, in order of misfit |
|------------------|------------|-----------|---------|---------|
| Item Number | Measure | SE | Infit MnSq | Outfit MnSq |
| 2 | -1.29 | 0.17 | 1.57* | 1.33 |
| 19 | -1.63 | 0.18 | 1.48* | 1.33 |
| 16 | -0.92 | 0.16 | 1.39 | 1.46* |
| 17 | 0.86 | 0.15 | 1.41* | 1.25 |
| 15 | 0.21 | 0.15 | 1.08 | 1.32 |
| 7 | 0.11 | 0.15 | 1.06 | 1.24 |
| 5 | 1.33 | 0.16 | 1.21 | 1.21 |
| 12 | -1.13 | 0.16 | 1.14 | 1.19 |
| 13 | -0.52 | 0.15 | 1.01 | 1.18 |
| 20 | 1.16 | 0.15 | 1.12 | 1.18 |
| 3 | 0.57 | 0.15 | 1.00 | 1.12 |
| 10 | -0.17 | 0.15 | 1.04 | 0.95 |
| 14 | -0.09 | 0.15 | 1.03 | 0.99 |
| 8 | 0.97 | 0.15 | 0.81 | 0.85 |
| 18 | 0.73 | 0.15 | 0.72 | 0.76 |
| 6 | 0.86 | 0.15 | 0.71 | 0.67 |
| 9 | -0.17 | 0.15 | 0.70 | 0.67 |
| 4 | -0.59 | 0.15 | 0.66 | 0.64 |
| 1 | 0.09 | 0.15 | 0.65 | 0.61 |
| 11 | -0.37 | 0.15 | 0.57* | 0.57* |

Each item estimate can be regarded as the balance point for the response distribution across that item’s categories. The higher the item estimate, the more difficult that item was for the group to endorse (higher scores). Items 3, 12, and 16 are reverse scored. Misfitting and overfitting values are marked by asterisks.
Despite the considerable item reduction from ABIS, the ABIS-R showed a large logit range in person ability (7.20 logits, from –3.21 to +3.99) and a reasonable logit range of item difficulty (2.70 logits, from –1.47 to +1.23) (Fig. 2B); this indicates a great spread of both person measures (as defined by the selected items) and item difficulties (as characterized by the sample under study). Av-

---

**FIGURE 2** Person-ability and item-difficulty maps of the ABIS (A) and ABIS-R (B). In each map, the double vertical line represents the measure of the variable, with the units of measurement on the scale (logits, the natural logarithm of the odds of mutually exclusive alternatives; e.g., pass vs. fail, or higher response vs. lower response). The lefthand column locates the individual person's performance along the variable: each person is indicated by an “X.” The righthand column locates the item difficulty measures along the variable (the difficulty estimate represents the mean calibration of the threshold parameters according to the rating-scale model, with values opposite those from Table 2). Each item is indicated by its number (see Appendix); fitting items are in bold. The top of the scale represents greater body image disturbance (patient measure) and greater item difficulty (higher item raw score). By convention, the average difficulty of items in the test is set at 0 logits (and indicated with $M$). Accordingly, a candidate with average ability is indicated with $M$. 

---

The table below illustrates the item difficulty and person ability measures for both ABIS and ABIS-R:

<table>
<thead>
<tr>
<th>ABIS</th>
<th>ABIS-R</th>
</tr>
</thead>
<tbody>
<tr>
<td>4++</td>
<td>4++</td>
</tr>
<tr>
<td>3++</td>
<td>3++</td>
</tr>
<tr>
<td>2++</td>
<td>2++</td>
</tr>
<tr>
<td>1+</td>
<td>1+</td>
</tr>
<tr>
<td>0+</td>
<td>0+</td>
</tr>
<tr>
<td>–1</td>
<td>–1</td>
</tr>
<tr>
<td>–3</td>
<td>–3</td>
</tr>
<tr>
<td>&lt;</td>
<td>&lt;</td>
</tr>
<tr>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>T</td>
<td>T</td>
</tr>
</tbody>
</table>

---

Gallagher et al.  
average person ability (M) was also at the same level
as average item difficulty (M'), which denotes that
the instrument was well matched to the sample
under study. Moreover, person abilities are nor-

TABLE 3 Correlation between the original Amputee Body Image Scale (ABIS) and revised
Amputee Body Image Scale (ABIS-R) and each of the Trinity Amputation Prosthesis
Experience Scales (TAPES) subscales

<table>
<thead>
<tr>
<th>Subscale</th>
<th>ABIS 20</th>
<th>ABIS-R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aesthetic satisfaction subscale (n = 119)</td>
<td>r = -0.27</td>
<td>R = -0.22</td>
</tr>
<tr>
<td>Weight satisfaction subscale (n = 121)</td>
<td>r = -0.23</td>
<td>R = -0.23</td>
</tr>
<tr>
<td>Functional satisfaction subscale (n = 120)</td>
<td>r = -0.41</td>
<td>R = -0.37</td>
</tr>
<tr>
<td>Athletic activity-restriction subscale (n = 114)</td>
<td>r = 0.17</td>
<td>R = 0.19</td>
</tr>
<tr>
<td>Functional activity-restriction subscale (n = 113)</td>
<td>r = 0.30</td>
<td>R = 0.31</td>
</tr>
<tr>
<td>Social activity-restriction subscale score less work (n = 120)</td>
<td>r = -0.40</td>
<td>R = -0.43</td>
</tr>
<tr>
<td>General adjustment subscale (n = 126)</td>
<td>r = -0.57</td>
<td>R = -0.54</td>
</tr>
<tr>
<td>Social adjustment subscale (n = 123)</td>
<td>r = -0.44</td>
<td>R = -0.40</td>
</tr>
<tr>
<td>Adjustment to limitations subscale (n = 116)</td>
<td>r = -0.30</td>
<td>R = -0.26</td>
</tr>
</tbody>
</table>

Finally, no floor or ceiling effects have been found.
In addition, the reliability values indicate that
ABIS-R has an acceptable internal consistency
(particularly for making group comparison) and a
high ability to define a distinct hierarchy of
items and persons along the measured construct.
Items were distributed into more than six difficulty
strata, and their high separation reliability (0.95)
indicates that high confidence can be placed in the
replicability of item placement across other sam-

dules. Furthermore, items of ABIS-R were able to
distinguish three levels of body image disturbances
(absent, low, high) in this sample of people with a
lower-limb amputation. Overall, these findings
show that ABIS-R presents a series of sound psy-
chometric properties.

The pattern of the item difficulty estimates in
the righthand columns of Figure 2B also provides a
description of some characteristics of feelings and

...I avoid situations where my physical appearance can be evaluated by others”) are more likely to
obtain higher ratings (i.e., answers in category 1
[sometimes] and 2 [more/all of the time]) than
items in the lower part of the column (e.g., items 5
and 20: “I avoid looking into a full-length mir-
or . . .”) or in the upper part but reverse scored
(e.g., item 12: “I like my physical appearance when
not wearing my prosthesis). Several characteristics
of this ordering are consistent with clinical expe-
rience and support the face validity of the item
hierarchy. For example, the physical appearance
when not wearing a prosthesis (item 12), the con-
cerns about the impairment of body’s functional
capabilities in various activities of daily living pro-
duced by a limb loss (item 4), and the worries that
“people notice my limp, when I’m walking” (item
13) are expected to be potential sources of distress
for people with lower-limb amputations. On the
contrary, the avoidance of looking into a full-
length mirror in order not to see one’s own pros-
thesis (item 5) or stump anatomy (item 20) is
expected to be infrequent in most people with a
lower-limb amputation. In addition, the two pairs
of similar items (items 5 and 20; items 10 and 14)
ocurred in the expected order of difficulty (Fig. 2B).

**Construct Validity**

The revised scale still contains items that re-
late to affective (e.g., item 1), cognitive (e.g., item
18), and behavioral (e.g., item 10) components of
body image as outlined by Breakey12 (See Ap-
pendix). The validity of the ABIS-R was also affirmed in
that similar relationships between the original ABIS and the ABIS-R with the TAPES subscales emerged despite the reduction in items. The relationships between the ABIS-R and each of the TAPES subscales were also in the hypothesized direction.

The moderate correlations between the ABIS-R and the “prosthesis satisfaction” subscales of the TAPES are not surprising and are consistent with the findings of Murray and Fox. As the prosthesis plays an important cosmetic and social role, in addition to a functional role, it is expected that a greater level of satisfaction with the prosthesis is associated with lower levels of body image disturbance. This association can be better appreciated if client satisfaction is considered as a concept related to the extent to which the features of prosthetic care received meet the client’s expectations, and if body image is considered as the portion of one’s self-concept that involves attitudes and experiences pertaining to the body that influence the individual’s subjective well-being. Thus, the personal background expectations and their perceived degree of fulfillment might be the shared underlying variable.

The association between ABIS-R and the activity-restriction subscales of TAPES was also in the expected direction; that is, lower levels of body image disturbance were associated with lower levels of activity restriction. This finding is consistent with Wetterhahn et al., who identified a positive relationship between regular participation in physical activity and body image among people with lower-limb amputations. Furthermore, Fisher and Hanspal recorded a significant correlation between body image and mobility in younger people with traumatic amputations. Indeed, Van Deusen postulated that body image distortion interferes with body movements that are necessary for daily activities to be performed. Finally, the correlation between the ABIS-R and the psychosocial adjustment subscales of the TAPES confirms the relationship between the body images of people with amputations and psychosocial adjustments to leg amputations found by Rybarczyk et al., Breakey, and Murray and Fox.

**CONCLUSION**

The 14-item ABIS-R demonstrates good psychometric characteristics for measuring body image disturbances in people with lower-limb amputations. Although these preliminary results suggest the adequacy of the new instrument, further studies are needed to analyze the actual performance of the new response structures and to confirm its measurement properties in other samples, thereby adding clinical validity to the instrument. Furthermore, a reasonable next step would be to assess whether ABIS-R items have different levels of difficulty (differential item functioning) on the basis of a sample characteristic such as gender or amputation level.

**APPENDIX**

Amputation Body Image Scale (ABIS). The 20-item questionnaire that, in the original version, was scored by a five-level ordinal scale (1 = none of the time, 2 = rarely, 3 = some of the time, 4 = most of the time, 5 = all of the time). Three of the questions (3, 12, and 16) are reverse scored. The 14 items of the shortened version (ABIS-R) are in bold, and their new scoring is 0 = none of the time, 1 = sometimes, and 2 = most/all of the time.

1. Because I am an amputee, I feel more anxious about my physical appearance in social situations than when I am alone.
2. I avoid wearing shorts in public because my prosthesis would be seen.
3. I like my overall physical appearance when wearing my prosthesis.
4. It concerns me that the loss of my limb impairs my body’s functional capabilities in various activities of daily living.
5. I avoid looking into a full-length mirror in order not to see my prosthesis.
6. Because I am an amputee, I feel anxious about my physical appearance on a daily basis.
7. I experience a phantom limb.
8. Since losing my limb, it bothers me that I no longer conform to society’s idea of normal appearance.
9. It concerns me that the loss of my limb impairs my ability to protect myself from harm.
10. When I am not wearing my prosthesis, I avoid situations where my physical appearance can be evaluated by others (e.g., I avoid social situations, swimming pool or beach activities, etc.).
11. The loss of my limb makes me think of myself as disabled.
12. I like my physical appearance when not wearing my prosthesis.
13. When I am walking, people notice my limp.
14. When I am wearing my prosthesis, I avoid situations where my physical appearance can be evaluated by others (e.g., I avoid any social situations, and/or I avoid swimming pool or beach activities etc.).
15. People treat me as disabled.
16. I like the appearance of my stump anatomy.
17. I wear baggy clothing in an attempt to hide my prosthesis.

1. Because I am an amputee, I feel more anxious about my physical appearance in social situations than when I am alone.
2. I avoid wearing shorts in public because my prosthesis would be seen.
3. I like my overall physical appearance when wearing my prosthesis.
4. It concerns me that the loss of my limb impairs my body’s functional capabilities in various activities of daily living.
5. I avoid looking into a full-length mirror in order not to see my prosthesis.
6. Because I am an amputee, I feel anxious about my physical appearance on a daily basis.
7. I experience a phantom limb.
8. Since losing my limb, it bothers me that I no longer conform to society’s idea of normal appearance.
9. It concerns me that the loss of my limb impairs my ability to protect myself from harm.
10. When I am not wearing my prosthesis, I avoid situations where my physical appearance can be evaluated by others (e.g., I avoid social situations, swimming pool or beach activities, etc.).
11. The loss of my limb makes me think of myself as disabled.
12. I like my physical appearance when not wearing my prosthesis.
13. When I am walking, people notice my limp.
14. When I am wearing my prosthesis, I avoid situations where my physical appearance can be evaluated by others (e.g., I avoid any social situations, and/or I avoid swimming pool or beach activities etc.).
15. People treat me as disabled.
16. I like the appearance of my stump anatomy.
17. I wear baggy clothing in an attempt to hide my prosthesis.
18. I feel I must have four normal limbs to be physically attractive.

19. It is important that my prosthesis and remaining anatomy of my affected limb are the same size as the other limb.

20. I avoid looking into a full-length mirror in order not to see my stump anatomy.

REFERENCES


16. Cash TF: The Multidimensional Body-Self Relations Questionnaire. Norfolk, VA, Old Dominion University, 1994


33. Wright BD, Linacre JM: Reasonable mean-square fit values. Rasch Meas Trans 1994;8:370

34. Wright BD, Masters GN: Rating Scale Analysis. Chicago, IL, MESA Press, 1982


38. McHorney CA, Monahan PO: Applications of Rasch analysis in health care. Med Care 2004;42 (1 Suppl): I73–8


Noninvasive Ventilation in Children with Spinal Muscular Atrophy Types 1 and 2

ABSTRACT


Objective: Our aim was to assess the efficacy of noninvasive ventilation (NIV) for the treatment of thoracoabdominal asynchrony during sleep in children with spinal muscular atrophy (SMA) types 1 and 2.

Design: Nine subjects underwent assessment for sleep apnea/hypopnea index (AHI), mean oxyhemoglobin saturation (SpO₂), oxygen desaturation index, transcutaneous carbon dioxide tension (tcpCO₂), and mean phase angle during sleep as a measure of thoracoabdominal coordination. A second sleep study was performed with use of NIV.

Results: The nine patients (7 mos of age, range 2–33) had a baseline AHI of 2.1 events per hour (range 0.5–55.8), oxygen desaturation index of 3.7 events per hour (range 1.6–46.1), mean tcpCO₂ of 46 mm Hg (range 37–60), and phase angle of 127 degrees (range 72.7–151.7). Comparing baseline and NIV sleep studies, we found significant improvement in oxygen desaturation index (P < 0.010), mean tcpCO₂ (P < 0.001), and phase angle (P < 0.001). For five patients, phase-angle improvement became significant when using high-span bilevel positive airway pressure (PAP).

Conclusions: NIV improved sleep breathing parameters and thoracoabdominal coordination during sleep in SMA types 1 and 2. Phase-angle improvement correlated with bilevel PAP pressures. Phase angle may be useful for the evaluation and monitoring of therapeutic interventions such as NIV.

Key Words: Spinal Muscular Atrophy, Respiratory Inductive Plethysmography, Noninvasive Ventilation, Thoracoabdominal Coordination, Sleep Apnea, Phase Angle
Spinal muscular atrophy (SMA) is an autosomal recessive congenital disease of the anterior horn cells of the spinal cord and the cranial nerves. It is characterized by weakness, with proximal muscles more affected; restrictive respiratory insufficiency; and absence of deep-tendon reflexes. This disorder is the leading hereditary cause of infant mortality.

SMA is stratified into five types based on severity of muscle weakness. SMA type 1 (Werdnig–Hoffmann disease) patients never attain the ability to sit independently; with conventional management, they usually die before age 2. Patients with SMA type 2 at least temporarily attain the ability to sit unsupported, but they also usually develop respiratory failure during childhood. Patients with SMA type 3 at least temporarily attain the ability to walk, and types 4 and 5 are adult onset. Use of respiratory muscle aids can greatly prolong survival for many patients with SMA types 1 and 2.

In patients with SMA, intercostal muscles are severely involved with relative sparing of the diaphragm. The intercostal muscle weakness is responsible for a funnel-shaped chest and pectus excavatum, paradoxical breathing, and diminished cough flows that result in atelectasis and bronchopulmonary infections. Respiratory muscle asynchrony increases the work of breathing and reduces tidal volume.

Paradoxical movements of the rib cage and abdomen are also seen physiologically during REM sleep in normal infants but decrease by age 3, and in otherwise normal children during obstructive sleep apneas. The SMA children have more asynchronous (paradoxical) breathing during active and quiet sleep than healthy controls and also when they are awake. In our previous study, we reported that both inspiratory and expiratory phases demonstrate thoracoabdominal dyssynchrony in these patients. Other earlier studies have included only older children and adults with a variety of neuromuscular conditions who were awake; these studies used predominantly invasive mechanical ventilation or low-span bilevel positive airway pressure (PAP). Perez et al. observed paradoxical breathing in 31 patients with SMA and 19 with myopathies with a mean age of 9.7 ± 3 yrs by using the percentage of thoracic contribution to tidal volume, a labored breath index, and phase angles. They also found that nearly full correction of chest-wall dyssynchrony was reported under these circumstances during intermittent positive pressure ventilation. Diaz et al. also observed chest-wall motion asynchrony in five children and young adults (mean age 15.6 yrs) with neuromuscular disease and its correction by noninvasive mechanical ventilation.

Mechanical ventilation can normalize blood CO₂ levels when hypercapnia is present, improve growth and development of lung parenchyma and the chest wall, rest inspiratory muscles, prevent pectus excavatum, prolong life, and improve quality of life. The aim of this study was to evaluate the utility of phase angle as a measure of the improvement of paradoxical chest and abdominal wall movement during sleep using noninvasive ventilation (NIV) delivered by nasal mask in infant patients with SMA types 1 and 2; these patients were much younger than those in the previous reports. Our goal was to observe the effects of varying bilevel PAP spans on achieve correction of dyssynchrony high-span nasal bilevel PAP and to study its effect on phase angle and other pulmonary parameters.

METHODS
Sample
We studied nine children with DNA deletion–confirmed SMA types 1 or 2, aged 7 mos (range 2–33), who had not had any signs of respiratory tract infection for at least 2 wks. Informed consent was obtained from the parents of each child. The ethics committee of our hospital approved this study.

Instruments and Technique
Sleep studies were performed on a respiratory ward and supervised by nurses. All studies were commenced at the patient’s usual bedtime and performed in a dark, quiet room. No sedation or sleep deprivation was used. All patients were accompanied by one of their parents throughout the night. A SomnoStar PT2 (Sensor Medics Corporation) was used to measure heart rate, pulse rate, pulse waveforms, arterial oxyhemoglobin saturation (SpO₂), calibrated respiratory inductive plethysmography (thoracic, abdominal, and summation channel), and oronasal airflow (thermistor). This technique has been used previously in children, and technical details have been published elsewhere. Transcutaneous carbon dioxide (tcpCO₂) was measured by a noninvasive monitoring system (TCM3, TINA, Radiometer; Copenhagen, Denmark). The instrument was calibrated before each study. CO₂ was monitored by transcutaneous methods for nocturnal CO₂ baseline and for bilevel PAP titration. No supplemental oxygen or respiratory stimulants were used.

Thoracoabdominal Coordination Parameters
The rib cage and abdominal waveforms define the two-compartment model of respiration described by Konno and Mead. These two components can move together or independently. Among the major indices that describe the coordination of movement between the rib cage and
abdominal compartments is the phase angle. A phase angle of 0 degrees indicates perfect in-phase movements, whereas values of 180 degrees indicate completely out-of-phase movements between the rib cage and abdominal compartments. Intermediate values indicate varying degrees of asynchronous and paradoxical motion. The phase angle represents the percentage agreement between direction of rib cage and abdominal movements through the entire cycle of a breath. The analysis of thoracoabdominal coordination was expressed as the mean of the entire sleep monitoring.

**Respiratory Parameter Definitions**

The polysomnogram is programmed to interpret apneas and hypopneas as being of obstructive or central origin rather than as a result of inspiratory muscle weakness. Because the polysomnogram cannot discern between these factors, we did not categorize the causes of the apneas and hypopneas.

The apnea/hypopnea index (AHI) was defined by the number of apnea/hypopneas per hour of sleep. For this study, sleep disordered breathing was defined as having an AHI greater than or equal to 1. The oxyhemoglobin desaturation index was defined as the number of desaturations of greater than or equal to 4% from baseline per hour of sleep. Oximetry artifacts in the pulse waveform tracing were discarded manually after visual inspection of the traces. Total sleep time was measured, and respiratory parameters were expressed as a percentage of total sleep time.

**Protocol Technique**

We prescribed high-span bilevel PAP (inspiratory minus expiratory bilevel PAP greater than 10 cm H2O) for SMA patients to promote more normal lung growth and chest-wall development. After the baseline polysomnogram was done, it was repeated using nasal bilevel PAP. It was introduced by using it for 20 mins, two or three times a day, before the sleep study. All patients used high-span bilevel PAP (range 14–20 cm H2O) to obtain good thoracic and alveolar expansion. Spans less than 10 cm H2O were considered low span.

Pressure preset ventilators cycling in pressure-assisted mode (A/C, ACHIEVA PSO2, Puritan Bennett-Tyco) were used. Nasal interfaces were selected to ensure a comfortable but airtight seal. Ventilator settings were chosen at bedside. To ameliorate patient–ventilator synchrony in patients unable to trigger the machine, the assisted/controlled modality of pressure-cycled machines was used with a back-up rate set slightly higher than the spontaneous breathing rate of the infant. As the increase of the tidal volumes resulted in a decrease in the infant’s spontaneous rate, the machine’s back-up rate was decreased and inspiratory PAP was augmented. Ventilator settings for bilevel PAP were initially set to achieve good excursion of chest wall and resolution of tachypnea and hypercapnia. The aim was to increase tidal ventilation, suppress patient respiratory trigger effort, decrease the patient’s spontaneous breathing rate, and optimize patient ventilator synchrony. Adjustments were made during the overnight sleep study to normalize gas exchange (mean SpO2 >95% and mean tcpCO2 <50 mm Hg) and to eliminate paradox.

**Statistical Analysis**

Statistical analysis was performed using the SPSS version 9.0.0 for Windows (SPSS, Inc., Chicago, IL). Data were expressed as median and range, except where otherwise specified. Baseline and bilevel PAP sleep study parameters were compared using the Mann–Whitney U test for nonparametric data. The threshold for significance was set at 0.05. Statistical applications for nonparametric data were used because of the small number of subjects studied.

**RESULTS**

**Subject Characteristics**

Subjects’ characteristics are shown in Table 1.

**Sleep Study Findings**

A summary of baseline sleep study findings is provided in Table 2. All subjects had mild to severe decreases in SpO2 and increases in thoracoabdominal asynchrony. A second cardiorespiratory sleep study was performed during bilevel PAP 10 days (range 5–14) after the baseline study, and the results are shown in Table 2.

There was a statistically significant improvement for oxygen desaturation index (P < 0.020), phase angle (P < 0.001), and tcpCO2 (P < 0.005) with the use of high-span bilevel PAP. No significant differences were found comparing the other parameters. Five patients performed an intermediate sleep study using a suboptimal bilevel PAP setting: inspiratory PAP 14 cm H2O (range 12–16), expiratory PAP 4 cm H2O (range 3–5) for a mean

| TABLE 1 Characteristics of children with spinal muscular atrophy (SMA) types 1 and 2 |
|---------------------------------|----|
| Age, mos                        | 7.0 (2–33) |
| SMA types 1                   | 4/5 (9) |
| Males/females                 | 7/2 |
| Weight, kg                     | 9.0 (4.4–16.0) |
| Height, cm                     | 95 (57–125) |

Unless otherwise specified, results are given as median (range).
span of 10 cm H2O (range 8–12). With this setting, the AH1 was reduced to no more than one event per hour (Table 3); however, the phase-angle improvement did not reach statistical significance: 110 degrees (range 90–145) vs. 88 degrees (range 54–110); P/H110210.080 (NS). The parents noted less nighttime perspiration and breathing difficulty during sleep using bilevel PAP.

### DISCUSSION

Overnight sleep studies performed during high-span bilevel PAP use by SMA type 1 and 2 patients showed an overall improvement in thoracoabdominal respiratory movements coordination as measured by phase angle, in mean tcpCO2, and in SpO2 desaturation rate.

All patients with SMA type 1 have paradoxical breathing. SMA type 1 and 2 patients who have paradoxical breathing develop pectus excavatum if untreated. This is caused by the action of the diaphragm that is not modulated by the intercostal muscles. In addition, when mechanical limitations of the ventilatory pump are superimposed on the physiologic reduction of compensatory inspiratory mechanisms during sleep with an increased upper-airway resistance, there can be serious impairment in a patient’s ability to maintain adequate respiration.23 This probably explains the nocturnal periods of perspiration and flushing and frequent arousals that have been reported to be relieved by the use of NIV.15

Bach et al.1 and Mellies et al.24 applied long-term NIV in children and adolescents with a variety of conditions. The results of these studies indicated that NIV can improve respiratory and sleep parameters in patients with SMA. For example, in one study, the mean AHI was reduced from 2.1 (0.5–55.8) to 0.6 (0.3–8.4) per hour, and the mean tcpCO2 was reduced from 46 (37–60) to 36 (35–38) mm Hg.

### TABLE 2

**Comparison between sleep study parameters during spontaneous breathing and noninvasive ventilation (NIV) in nine children with spinal muscular atrophy types 1 and 2**

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Baseline Sleep Study (n = 9)</th>
<th>High-Span NIV Sleep Study (n = 9)</th>
<th>P Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>TST, hrs</td>
<td>8.2 (7.0–9.4)</td>
<td>8.0 (7.0–8.9)</td>
<td>0.851 (NS)</td>
</tr>
<tr>
<td>Mean SaO2, %</td>
<td>96 (92–97.9)</td>
<td>97 (95–98)</td>
<td>0.373 (NS)</td>
</tr>
<tr>
<td>SpO2 &lt;90%, % TST</td>
<td>0.1 (0.0–36.1)</td>
<td>0.0 (0.0–6.7)</td>
<td>0.244 (NS)</td>
</tr>
<tr>
<td>SpO2 &lt;95%, % TST</td>
<td>3.4 (1.6–46.1)</td>
<td>0.3 (0.0–90.1)</td>
<td>0.184 (NS)</td>
</tr>
<tr>
<td>ODI, n ≥4%/hr</td>
<td>3.0 (1.6–46.1)</td>
<td>0.9 (0.0–22.0)</td>
<td>&lt;0.020</td>
</tr>
<tr>
<td>AH1, n/hr</td>
<td>2.1 (0.5–55.8)</td>
<td>0.6 (0.3–8.4)</td>
<td>0.101 (NS)</td>
</tr>
<tr>
<td>tcpCO2 mean, mm Hg</td>
<td>46 (37–60)</td>
<td>36 (35–38)</td>
<td>&lt;0.005</td>
</tr>
<tr>
<td>Mean phase angle, degrees</td>
<td>127 (72.7–151.7)</td>
<td>44 (38–68)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Results are given as median (range).

The high-span NIV sleep study parameters were obtained using the following settings: inspiratory PAP 18 cm H2O (range 14–20); expiratory PAP 4 cm H2O (range 3–5); bilevel PAP span 14 cm H2O (range 10–15). The baseline sleep study was performed during spontaneous breathing. The high-span NIV sleep study was performed using bilevel PAP. TST, total sleep time; mean SpO2, mean arterial oxyhemoglobin saturation; SpO2 <90% (%TST), percentage of TST spent with an oxyhemoglobin saturation (SpO2) below 90%; SpO2 <95% (%TST), percentage of TST spent with an oxyhemoglobin saturation (SpO2) below 95%; ODI, oxyhemoglobin desaturation index; AH1, apnea/hypopnea index; tcpCO2, transcutaneous CO2.

### TABLE 3

**Comparison between sleep study parameters during noninvasive ventilation (NIV) using different bilevel PAP spans in five children with spinal muscular atrophy types 1 and 2**

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Intermediate NIV Sleep Study (n = 5)</th>
<th>High-Span NIV Sleep Study (n = 5)</th>
<th>P Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bilevel PAP span, cm H2O</td>
<td>10 (8–12)</td>
<td>14 (10–15)</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>TST, hrs</td>
<td>8.2 (7.9–8.6)</td>
<td>7.9 (7.0–8.7)</td>
<td>0.341 (NS)</td>
</tr>
<tr>
<td>Mean SpO2, %</td>
<td>96.7 (95.8–98)</td>
<td>97 (95–97.7)</td>
<td>0.917 (NS)</td>
</tr>
<tr>
<td>SpO2 &lt;90%, % TST</td>
<td>0.0 (0.0–0.2)</td>
<td>0.0 (0.0–0.0)</td>
<td>0.136 (NS)</td>
</tr>
<tr>
<td>SpO2 &lt;95%, % TST</td>
<td>0.5 (0.2–4.2)</td>
<td>0.2 (0.0–1.2)</td>
<td>0.115 (NS)</td>
</tr>
<tr>
<td>ODI, n ≥4%/hr</td>
<td>2.0 (1.1–2.3)</td>
<td>0.1 (0.0–1.0)</td>
<td>0.101 (NS)</td>
</tr>
<tr>
<td>AH1, n/hr</td>
<td>0.8 (0.5–1.0)</td>
<td>0.6 (0.3–0.8)</td>
<td>0.242 (NS)</td>
</tr>
<tr>
<td>Mean tcpCO2, mm Hg</td>
<td>38 (32–40)</td>
<td>36 (35–38)</td>
<td>0.338 (NS)</td>
</tr>
<tr>
<td>Mean phase angle, degrees</td>
<td>88 (54–110)</td>
<td>48 (38–68)</td>
<td>&lt;0.03</td>
</tr>
</tbody>
</table>

Results are given as median (range).

The high-span NIV sleep study was performed using high-span bilevel PAP. Bilevel PAP span, inspiratory PAP–expiratory PAP expressed as centimeters of water. TST, total sleep time; mean SpO2, mean arterial oxyhemoglobin saturation; SpO2 <90% (%TST), percentage of TST spent with an oxyhemoglobin saturation (SpO2) below 90%; SpO2 <95% (%TST), percentage of TST spent with an oxyhemoglobin saturation (SpO2) below 95%; ODI, oxyhemoglobin desaturation index; AH1, apnea/hypopnea index; tcpCO2, transcutaneous CO2.
of congenital neuromuscular disorders (intermediate SMA, congenital muscular dystrophy and myopathies) who presented with either nocturnal hypoventilation or diurnal ventilatory failure. They reported similar improvements in sleep quality and symptoms. Eighty-one percent of the neuromuscular patients reported in another study were maintained on nocturnal ventilation for years. The authors postulated that these children remained stable because the nocturnal ventilation ensured adequate gas exchange during sleep, reversed atelectasis, decreased inspiratory muscle work, and rested the respiratory muscle.25,26 Metabolic abnormalities caused by chronic hypoventilation were also ameliorated by the resetting of central chemoreceptors and/or improving sleep fragmentation.23,27 Barois et al.28 advocated undertaking ventilator use before the age of 4 yrs in SMA type 2 patients to improve rib-cage development and foster lung growth. It has been further suggested that earlier intervention with respiratory support may actually allow for greater lung growth and improved respiratory outcome.15

Previous studies13,14 have shown improvement in thoracoabdominal asynchrony by measuring respiratory inductive plethysmography in older children and adults using mechanical ventilation. Perez et al.13 studied 31 patients with SMA and 19 patients with myopathies, aged 9.7 ± 3 yrs. All patients were studied in the supine position for several minutes, both during spontaneous breathing and mechanical ventilation, apparently while awake. They found that percent thoracic contribution to tidal volume, labored breath index, and phase angle were normalized during intermittent positive pressure ventilation. Diaz et al.14 studied five patients aged 15.6 yrs (range 9–23) with different neuromuscular conditions including C-2, C-3, and T-4 transaction, congenital myopathy, and SMA in the supine position and apparently awake. Chest-wall motion was recorded for several minutes in spontaneous breathing and during mechanical ventilation. Mechanical ventilation was performed using negative pressure ventilation in one subject, positive pressure nasal ventilation in another, and volume-cycled positive pressure ventilation via tracheostomy tube in the three remaining subjects. Diaz et al.14 considered percentage of thoracic contribution to tidal volume, a labored breath index, and phase angles and concluded that mechanical ventilation improves thoracoabdominal asynchrony and reverses the negative contribution to tidal volume of the paradoxing compartment in children and young adults with neuromuscular disease. Neither study correlated changes with ventilator settings.

Our study concerned only sleeping infants with SMA types 1 and 2, using only NIV, and demonstrated improvement in thoracoabdominal coordination and respiratory movements that correlated with bilevel PAP spans. Our data suggest that not only can nocturnal NIV normalize AHI and SpO2 when used at low bilevel PAP spans, but also that at higher spans, it can significantly normalize inspiratory muscle synchrony. This can explain the previously reported prevention of pectus excavatum.29,30 Results of this study indicate that NIV can improve alveolar ventilation and thoracoabdominal respiratory movement coordination.

In conclusion, NIV improved sleep breathing parameters and thoracoabdominal coordination during sleep in SMA types 1 and 2. Phase-angle improvement correlated with bilevel PAP pressures. Phase angle may be useful for the evaluation and monitoring of therapeutic interventions such as NIV.

REFERENCES

15. Bach JR, Bianchi C: Prevention of pectus excavatum for...
Avoiding Respiratory Failure in Neuromuscular Disease
Why Is It Not Done?

ABSTRACT

Key Words: Noninvasive Mechanical Ventilation, Neuromuscular Disease, Spinal Muscular Atrophy, Werdnig–Hoffmann Disease, Duchenne Muscular Dystrophy, Amyotrophic Lateral Sclerosis

The severity of spinal muscular atrophy (SMA) type 1 (Werdnig–Hoffmann disease) can be characterized as severe when the infant definitively requires continuous ventilatory support before 5 mos of age, moderate when the child first develops acute respiratory failure as the result of an ineffective cough during an acute respiratory tract infection from 5 to 24 mos of age, and mild when the child does not develop respiratory failure and require a gastrostomy for full nutrition until after the second birthday.1 Considering only the severely and moderately affected children, death by 2 yrs of age is unavoidable unless the infant undergoes tracheotomy. In this case, the child remains continuously ventilator dependent, does not develop functional speech, and most likely dies from complications related to the tracheostomy tube.2 However, when these children develop acute respiratory failure and the acute infection has cleared, they can be extubated to noninvasive mechanical ventilation (NIV) regardless of whether they have any ability to breathe on their own.1,2 Likewise, it has been demonstrated that by intermittently applying negative pressure to the airway to support expiratory (cough) muscles and by using concomitant abdominal thrusts (mechanically assisted coughing, or MAC), pneumonias and hospitalizations can be avoided in SMA 1 patients using NIV, without need to resort to tracheotomy.1,2

In addition to prolonging life by using NIV continuously if necessary, the introduction of nocturnal NIV for all hypotonic children with paradoxical breathing promotes lung and chest-wall growth and prevents pectus excavatum.3 Despite this, third-party payers often mandate abnormal polysomnography results before paying for NIV, even though the issue is not central or obstructive apneas but weak inspiratory muscles that require rest and assistance.

When normal babies cry, their abdomens protrude and their chest walls expand. This promotes lung and chest-wall growth. However, the crying of a child with paradoxical breathing only exacerbates paradox, and their lungs and chest walls do not grow normally. The minimum pressure required to provide...
a normal tidal volume and expand the chest is 18–20 cm H$_2$O for normally compliant lungs and chest walls at any age. Effective inspiratory muscle rest also requires NIV at these inspiratory pressures. Thus, even if polysomnography is “abnormal” and apneas and hypopneas are noted, continuous positive airway pressure will not assist inspiratory effort, and titration of bilevel positive airway pressure (PAP) to reduce hypopneas is suboptimal. This is because any pressures less than 15–20 cm H$_2$O will simply not well rest inspiratory muscles, high-span bilevel PAP or NIV at full ventilatory-support volumes and pressures is indicated.

Because it is clear that children with SMA, muscular dystrophy, patients with nonbulbar amyotrophic lateral sclerosis, postpolio, high-level traumatic tetraplegia, and other conditions can, and prefer to, use NIV and MAC rather than simply resorting to tracheotomy, the full cooperation of intensive care physicians is needed to extubate these patients by specifically described protocols to NIV and MAC, rather than simply resorting to tracheotomy. The almost continuous attention to maintain normal SpO$_2$ (>94% in ambient air) by using oximetry as feedback for effective MAC, especially after extubation, can necessitate continuous and active participation of a family member or care provider, because hospital staff often do not have the time to provide this. Allowing unlicensed care providers to take part in intensive care also requires a new treatment paradigm, but it is critical for successful extubation, rehabilitation, and transition of care to the home.

In addition, neither the primary care nor intensive care physician can manage these patients without specifically trained respiratory therapists for both home and intensive care. Our Jerry Lewis Muscular Dystrophy Association Clinic grant provides funds for a respiratory therapist to evaluate and train patients in the clinic in these methods. This is preferable to using the funds for more expensive pulmonary function testing that was neither designed for, nor adequate for, patients with muscle impairment rather than lung/airways disease.

—Kenneth Olsen, 1977 president and founder of Digital Equipment Corporation

—[Television] won’t be able to hold on to any market it captures after the first six months. People will soon get tired of staring at a plywood box every night.”

—1946, Darryl F. Zanuck, head of 20th Century-Fox

—The earth is the center of the universe.”

—The first-century Ptolemy of Alexandria, and the Popes of the Roman Catholic Church

To this can now be added:

• Determining the best date for tracheostomy in patients with (neuromuscular disease) remains a challenge.”


The need for an NIV team: Besides the primary care physician who needs to introduce NIV and assisted coughing to prevent acute respiratory failure, when such episodes occur despite these efforts, the full cooperation of intensive care physicians is needed to extubate these patients by specifically described protocols to NIV and MAC, rather than simply resorting to tracheotomy. The almost continuous attention to maintain normal SpO$_2$ (>94% in ambient air) by using oximetry as feedback for effective MAC, especially after extubation, can necessitate continuous and active participation of a family member or care provider, because hospital staff often do not have the time to provide this. Allowing unlicensed care providers to take part in intensive care also requires a new treatment paradigm, but it is critical for successful extubation, rehabilitation, and transition of care to the home.

In addition, neither the primary care nor intensive care physician can manage these patients without specifically trained respiratory therapists for both home and intensive care. Our Jerry Lewis Muscular Dystrophy Association Clinic grant provides funds for a respiratory therapist to evaluate and train patients in the clinic in these methods. This is preferable to using the funds for more expensive pulmonary function testing that was neither designed for, nor adequate for, patients with muscle impairment rather than lung/airways disease.

—Kenneth Olsen, 1977 president and founder of Digital Equipment Corporation

—[Television] won’t be able to hold on to any market it captures after the first six months. People will soon get tired of staring at a plywood box every night.”

—1946, Darryl F. Zanuck, head of 20th Century-Fox

—The earth is the center of the universe.”

—The first-century Ptolemy of Alexandria, and the Popes of the Roman Catholic Church

To this can now be added:

• Determining the best date for tracheostomy in patients with (neuromuscular disease) remains a challenge.”


The need for an NIV team: Besides the primary care physician who needs to introduce NIV and assisted coughing to prevent acute respiratory failure, when such episodes occur despite these efforts, the full cooperation of intensive care physicians is needed to extubate these patients by specifically described protocols to NIV and MAC, rather than simply resorting to tracheotomy. The almost continuous attention to maintain normal SpO$_2$ (>94% in ambient air) by using oximetry as feedback for effective MAC, especially after extubation, can necessitate continuous and active participation of a family member or care provider, because hospital staff often do not have the time to provide this. Allowing unlicensed care providers to take part in intensive care also requires a new treatment paradigm, but it is critical for successful extubation, rehabilitation, and transition of care to the home.

In addition, neither the primary care nor intensive care physician can manage these patients without specifically trained respiratory therapists for both home and intensive care. Our Jerry Lewis Muscular Dystrophy Association Clinic grant provides funds for a respiratory therapist to evaluate and train patients in the clinic in these methods. This is preferable to using the funds for more expensive pulmonary function testing that was neither designed for, nor adequate for, patients with muscle impairment rather than lung/airways disease.

—Kenneth Olsen, 1977 president and founder of Digital Equipment Corporation

—[Television] won’t be able to hold on to any market it captures after the first six months. People will soon get tired of staring at a plywood box every night.”

—1946, Darryl F. Zanuck, head of 20th Century-Fox

—The earth is the center of the universe.”

—The first-century Ptolemy of Alexandria, and the Popes of the Roman Catholic Church

To this can now be added:

• Determining the best date for tracheostomy in patients with (neuromuscular disease) remains a challenge.”

trists, and few physiatrists are interested in respiratory muscle aids.16

“A pulmonologist who considers the airways but who ignores the respiratory muscles is like a cardiologist who considers the blood vessels but who ignores the heart.”
—Peter Maclem, MD

Money: Physicians are paid more for placing invasive tubes than for avoiding them. Hospitals and physicians are compensated for the intensive care management of patients in respiratory failure, but preventing respiratory failure with appropriate home care requires time, effort, and letters of explanation to third-party payers to provide the needed training and equipment.

“No matter what you say or how, nothing talks like money!”
—Petronius Arbiter, the Satyricon (time of Nero)

Assistance: Patients who are managed noninvasively require continuous attention for MAC during intercurrent respiratory tract infections to avoid pneumonia and acute respiratory failure. This must be provided by families or personal care attendants, to whom not all patients have sufficient access.

Inadequate Equipment: In a survey of 88 hospitals, it was found that almost none were adequately equipped (absence of portable noninvasive ventilators, noninvasive interfaces, mechanical in-exsufflators) and that staff were inadequately trained for NIV and assisted coughing.17

Thus, although tracheostomy tubes are avoidable for ventilator-dependent infants with neuromuscular disease and are even easier to avoid for older children and adult ventilator users, until physiatrists become actively involved in the new paradigms of evaluation and management, most patients will continue to suffer potentially avoidable episodes of acute respiratory failure and die prematurely or live with foreign bodies in their necks. The latter renders home care more difficult, or even impossible to achieve, and results in an unnecessarily poorer quality of life and greater ventilator dependence.2,6,14

REFERENCES
Introduction to Nanotechnology
Potential Applications in Physical Medicine and Rehabilitation

ABSTRACT

Nanotechnology is a scientific movement that has the potential to transform the diagnosis and treatment of disease in the 21st century. The area of investigation is defined by the study, design, manipulation, manufacture, and control of materials or devices by physical or chemical means at resolutions on the order of one billionth of a meter. The potential for a wide range of clinical applications makes a basic understanding of nanotechnology important to physiatrists. This review presents an introduction to nanotechnology and discusses key developments in tissue engineering, drug delivery, imaging, diagnostics, surface texturing, and biointerfaces that could impact the practice of physiatry in the future.

Key Words: Nanotechnology, Rehabilitation, Tissue Engineering, Drug Delivery, Imaging, Diagnostic Assays

The broad applications of nanotechnology make it one of the most rapidly growing areas of research in contemporary science and medicine. The working definition of nanotechnology involves research on materials and systems whose sizes range between 1 and 100 nanometers (nm)—approximately the scale of the basic building blocks of matter and molecules (Fig. 1). The field is not a scientific discipline unto itself but, rather, a confluence of chemistry, physics, engineering, and material science. Much attention has been recently focused by biomedical scientists, pharmaceutical companies, and investors on the potential impact of this rapidly expanding field to medicine. Worldwide public investment in all uses of nanotechnology was $8.6 billion in 2004, and in 2005 the U.S. government appropriated nearly $1 billion in federal funding among 11 federal agencies.1 Approximately 90 million dollars were appropriated in 2005 for biomedical research and development to the U.S. National Institutes of Health through the National Nanotechnology Initiative. In March 2005, the first volume of the quarterly journal Nanomedicine was published, an important addition to the multiple technical nanotechnology journals in nonmedical fields.

In both 2004 and 2005, the U.S. National Advisory Board on Medical Rehabilitation Research identified nanotechnology as a research opportunity area of focus, with an identified shortage of faculty with training in the rehabilitative
specialties and a need for clinical trials. Several priority areas in nanoscience have been identified by the U.S. National Institutes of Health Bioengineering Consortium, including tissue engineering, imaging, cell signaling research, implantable bioproces- sors, biocompatible materials, drug delivery agents, and sensor technologies.2

In combination with advancements in biochemistry and engineering, nanotechnology has shown promise for future clinical applications across varied fields. This review of nanotechnology will focus on the potential applications to rehabilitation medicine. It is intended to be introductory, not an in-depth critical analysis of the many technical complexities and limitations in this rapidly growing field. Given this framework, the goals of this review are to (1) introduce nanotechnology to the rehabilitation community, (2) outline a brief history of nanotechnology, (3) focus discussion of nanotechnology to those specific areas of research that will impact the field of physical medicine and rehabilitation, and (4) encourage interest among physiatrists in this growing area of scientific investigation. The reader is directed to (Appendix A for definitions of possibly unfamiliar nanotechnology terms (first appearance in italics) in this review.

BRIEF HISTORY OF NANOTECHNOLOGY

In December 1959, Richard Feynman, a theoretical physicist well known for his work on the Manhattan Project and the development of the atomic bomb, delivered a lecture entitled “There’s Plenty of Room at the Bottom.” This often-quoted speech has become known as the landmark prediction of the potential field of nanotechnology. Dr. Feynman conceived of a time when visualization, manipulation, data storage, and manufacture would be possible in the atomic size range. He asked, “Why cannot we write the entire 24 volumes of the Encyclopedia Britannica on the head of a pin?”

Many of the specific aspects of the field addressed in the lecture required 40 yrs of work before they could indeed be realized or understood. Dr. Feynman anticipated several hurdles to the development of the field, including increased resistance of circuits on a small scale and overheating of nanoprocessors. He also foresaw a nanoworld dominated by different physical properties such as van der Waals forces and Brownian motion that would only come to be understood after extensive and prolonged study.

The field remained largely theoretical until the early 1980s, when several key advancements demonstrated the successful practical application of nanotechnology. In 1981, Gerd Binnig and Heinrich Rohrer (who in 1986 shared the Nobel Prize in Physics) developed the scanning tunneling microscope, which allowed scientists to see substances on the level of individual atoms.3 Also in the early 1980s, Eric Drexler, then a Massachusetts Institute of Technology graduate student, developed a thesis that later became a pioneering text entitled “Engines of Creation: The Coming Era of Nanotechnology,” published in 1985.4 Drexler foresaw the bottom-up design of proteins that would fabricate complex nanodevices with customized specifications. These nanomachines would mimic biological processes and be capable of self-replication, specialized function, and even self-destruction. Also in 1985, Richard Smalley, Robert Curl, and Sir Harold Kroto5 discovered the molecular structures of fullerenes (for which they later were awarded the Nobel Prize), and a team of IBM scientists manipulated 35 xenon molecules on a gold surface to spell out the letters “IBM.” Twenty-six years after Feynman’s prediction, nanowriting had become a reality.

Nanomedicine

Although research in nanotechnology has been broadly expanding in many areas of science during the past 15 yrs, it has been only recently that its applications to medicine have undergone rapid development (Fig. 2).
Appendix A  Glossary of nanotechnology definitions related to biosciences

**Bottom-up vs. top-down**  
“Bottom-up” is the term used to describe starting with single atoms and molecules to build nanostructures. This concept describes technologies enabling the construction of things atom by atom to achieve molecular-level structures. The phrase “top-down” refers to methods and tools to build small components starting from larger regular matter.

**Brownian motion**  
The physical phenomenon that minute particles immersed in a fluid move around randomly, or the mathematical model used to describe those random movements. It was discovered by botanist Robert Brown in 1827 while studying pollen particles floating in water under the microscope.

**Buckyball**  
Short for buckminsterfullerene, a molecule of carbon made of 60 atoms and shaped like a geodesic dome. The buckyball was discovered by Richard Smalley and colleagues at Rice University in 1985. Smalley received the Nobel Prize for this discovery. Buckyballs are inert, nontoxic, and infinitely modifiable to form new formulations. The molecule is hollow and has been investigated as a transporter of medicines and contrast. Buckyballs are the third form of pure carbon known to exist naturally, along with graphite and diamonds. The process of knocking one or more carbon atoms out of the Buckyball structure and replacing it with metal atoms is known as doping, and the molecule in its altered form is referred to as a dopeyball.

**Dendrimer**  
A synthetic, tree-shaped macromolecule formed from monomers using a stepwise fabrication process. It is the size of a typical protein, but it does not unfold or come apart easily because it is held together with stronger chemical bonds. It can be easily attached to a variety of other molecules to transport and release medication/contrast or to transport DNA for gene therapy.

**Electrospinning**  
A fabrication process that uses an electric field to control the deposition of polymer fibers onto a target tissue. During the process, a polymer solution is injected with an electrical potential onto a grounded target. The solvent then evaporates and fibers are formed. This can be used to fabricate fibrous polymer scaffolding for use in tissue engineering.

**Fullerines**  
A class of carbon molecules, of which buckyballs are a member. Fullerenes are closed, convex cage molecules containing only hexagonal and pentagonal faces. These molecular forms of pure carbon are made when vaporized carbon condenses in an atmosphere of inert gas. Although C60 is the most common, there are larger fullerenes containing 70–500 carbon atoms.

**Gene chip**  
DNA microarrays used to screen a biological sample for the presence of many genetic sequences with affixed DNA segments known as probes. This technology is being evaluated as a replacement for the current costly and time-consuming polymerase chain reaction methods. By spotting thousands of DNA pieces on a small surface (microchip), thousands of DNA experiments can be completed in a single run, as opposed to one at a time by polymerase chain reaction.

**Lab on a chip**  
A microfluidic device that is a promising analytic tool for analyzing proteins and protein complexes. These “laboratories” are fabricated using photolithographic processes developed in the microelectronics industry to create circuits of tiny chambers and channels in a quartz, silica, or glass chip. They direct the flow of liquid chemical reagents just as semiconductors direct the flow of electrons. These microfluidic circuits can be designed to accommodate virtually any analytic biochemical process and allow for parallel chemical analyses.

**Lithography**  
A printing process used in fabrication, often in the electronics industry. There are several types, including a) photolithography, in which an electron beam writes the circuit pattern on a microchip; b) soft lithography, in which an elastic stamp produces patterns that hold a solidifying polymer; and c) dip-pen lithography, in which an AFM coats a gold surface with a self-assembled monolayer of thiol molecules. Several other processes are being investigated that can either adapt existing microtechnology to current electronic production or that would require entirely new production techniques and investment for electronic circuit manufacture.

**Nanofibers**  
Structures that do not possess as regular a helical carbon arrangement as carbon nanotubes and, thus, have slightly less impressive mechanical characteristics. They are also far cheaper to produce than nanotubes.

**Nanometer (nm)**  
Unit of measurement: 1 billionth of a meter. 1 nm = 10 angstrom units (one angstrom is the diameter of a hydrogen atom—the smallest element). 1000 nm = one micron. To give a sense of proportion: hydrogen atom = 0.1 nm, protein = 5–50 nm, gene = 2–5 nm wide and 10–100 nm long, virus = 20–450 nm, and red blood cell = 10,000 nm wide.

**Nanoparticle**  
A microscopic particle whose size is measured in nanometers. Often, such nanoscale particles are used in biomedical applications as drug carriers or imaging agents.

**Nanopores**  
Tiny holes that are used in filters, sensors, or tissue scaffolds that enhance function. For instance, nanopores allow DNA to pass through one strand at a time, which makes DNA sequencing more efficient. Nanopores can also span cell membranes and allow ionic transport across the otherwise impermeable lipid bilayer.
### Appendix A Continued

| Nanoshell | A 100-nm bead composed of a spherically shaped metal (often gold) surrounding a core of silicon dioxide atoms. Because of their size, nanoshells can easily penetrate several centimeters of tissue and can be functionalized by linking to antibodies. They are actively being pursued as a treatment for certain cancers because they capture energy at near infrared and deform, releasing a chemotherapeutic drug at the tumor site. |
| Nanotube (NT) | Discovered in 1991 by Japanese researcher Sumio Iijima, a nanotube is a sequence of nanoscale C60 atoms arranged in a long, thin, cylindrical structure. Carbon nanotubes have many potential applications because of their specific properties; they are 100 times stronger than steel (and six times lighter), they are as conductive as copper, and they can resist very high temperatures (up to 1500°C under vacuum). Applications of the nanotubes include resistors, capacitors, inductors, diodes, and transistors. The discovery of nanotubes has initiated an entirely new field of scientific research. A multi-wall carbon nanotube (MWNT) is a cylindrical piling up of single-wall carbon nanotubes (SWNT). |
| Nanowire | A one-dimensional structure with unique electrical or optical properties. Easy to produce, nanowires can be assembled in grids and become the basis of nanoscale logic circuits. They can have very different shapes. Often thin and short “threads,” they can look like cones, carrots, cherries, or comets. Latticed in various combinations, nanowires can function as transistors, optoelectronic devices, and biochemical sensors. |
| National Nanotechnology Initiative (NNI) | Established in 2001, a federal program to support research and development of the nanosciences in the United States. The initiative is designed to develop the necessary infrastructure to support education, facilities, instrumentation, and communication in both the public and private sectors. Its vision is to achieve international leadership in nanotechnology and technology to promote sustainable economic benefit and national security. Over 20 government agencies are either funded or participate in the NNI. In the first year, funding appropriations totaled $464 million. Currently, investment totals are near $1 billion. |
| Phase separation | A fabrication method to construct synthetic polymer nanofiber matrices in which a polymer solution is separated while in contact with a structured mold. Intrinsic shrinkage during the process facilitates the release of the replica from a mold. The unique advantage of this technique is the capability of designing three-dimensional nanostructures without using any special equipment such as clean rooms. Also, the porosity of a polymer scaffold can be easily modified and controlled. |
| Quantum dots (QDs) | Nanometer-sized semiconductor crystals that are 5–10 nm in size and that can be dissolved in water. Their cores contain paired clusters of atoms such as cadmium and selenium that combine to create a semiconductor. Ultraviolet light of wide range can be used to stimulate these crystals, but they illuminate at only one specific wavelength, depending on the size and number of atoms in the core. When biological molecules are attached to the quantum dots, this specificity of illumination allows them to be used as probes to track antibodies, viruses, proteins, or DNA. The size of the dot also determines its magnetic, electronic, and structural properties. |
| Scanning tunneling microscope (STM) | An instrument that probes surface contours. It detects small currents between the sharp stylus mounted on a soft spring and the sample, allowing imaging on the scale of individual atoms. Critical to the nanotechnology revolution, the STM was invented in 1981 by IBM’s Gerd Binnig and Heinrich Rohrer, who, in 1986, shared the Nobel Prize in Physics for this discovery. |
| Self-assembly | Self-assembly refers to the joining of complementary surfaces, leading molecules to arrange themselves automatically into higher-order structures. These molecules follow a regular pattern under certain pressure and temperature conditions. Self-assembly allows production of nanostructures or nanopatterns with interesting properties, in large quantities, with low effort and total control of the reaction. |
| Ultrasound particles of iron oxide (USPIO) | Magnetic nanoparticles used in enhancing magnetic resonance imaging of metastatic tissue and in staining of target cells and molecular processes. Similar entities include supermagnetic nanoparticles and cross-linked iron oxide nanoparticles. |
| Van der Waal forces | The electrical intermolecular attractions between one molecule and a neighboring molecule (after botanist Johannes Diderik van der Waals, 1837–1923). Temporary fluctuating dipoles form dispersion forces that attract particles together (also known as London forces). These forces become stronger as the atom (or molecule) becomes larger and thinner. |

Some general advances in nanotechnology have indirectly impacted biomedicine. For instance, new methods for developing computer nanocircuits have allowed for smaller and more effective medical appliances (e.g., retinal prostheses, implantable microdevices for neural control of movement). Although the indirect and “spin-off” technology may be noteworthy, this review focuses on the directly applied nanotechnology in biomedical research and clinical practice across
five major areas: (1) tissue engineering, (2) drug and gene delivery, (3) imaging, (4) diagnostic assays, and (5) texturing/biointerfaces. A brief general background will be provided in each area, followed by a more focused review of research specifically related to the field of rehabilitation medicine.

**Tissue Engineering**

The basic principles of tissue engineering and regeneration include in vitro and in vivo generation of tissue. For the in vitro process, the tissue construct begins with specialized cells (e.g., stem cells\(^8,9\)) placed in an environment conducive to cell growth, often in an artificial extracellular matrix or scaffolding. The extracellular matrix was once thought to play largely a structural role. However, it now is recognized to play a major role in cell proliferation, differentiation, and inhibition of dedifferentiation.

Nanotechnology has been used largely to investigate improvements in the fabrication techniques of synthetic or organic scaffolding. Controlled design allows for manipulation of structural cues, cell attachment, and intercellular signaling.\(^10\) Fibrous meshes, porous sponges, or hydrogels can be modified to mimic natural matrix structure, embedded with growth-specific surface proteins (to promote such factors as angiogenesis\(^11\)), and designed to biodegrade in a controlled fashion (to allow for overgrowth by innate cells). The most popular techniques of fabrication of polymer nanofibers into scaffolding are *electrospinning*, *phase separation*, and *self-assembly*.\(^12\)

The active research in regeneration of damaged brain, spinal cord, skin, cartilage, muscle, and bone tissues is likely to impact the future of rehabilitation medicine. A compelling and highly publicized series of experiments on the potential benefits of nanoscaffolding in brain repair and axonal regeneration was conducted by Ellis-Behnke et al.\(^13\) In this animal model of brain damage, the optic tract within the midbrain was completely severed with a deep knife wound in young and adult hamsters. In one experiment, ten animals in the treatment group received an injection into the wound of a self-assembling peptide that spontaneously formed a scaffolding network of nanofibers. The remaining six received injections of isotonic saline. Both groups were subsequently analyzed histologically for healing of injury and behaviorally for functional restoration of vision.

In all animals that received self-assembling peptide, the tissue gap created during injury was either reduced or completely eliminated at time points greater than 24 hrs after injury. All self-assembling peptide–treated animals showed a significant percentage of axonal reinnervation as measured by fluorescent staining, in contrast to no axonal regeneration in the control group. Most importantly, the self-assembling peptide–treated hamsters gained an average of 75% recovery of visual ability as indicated by the ability to orient to a small object. The control group remained blind. This model was successful in demonstrating that nanoscaffolding can create an environment conducive to functional axonal regrowth in brain tissue while reducing scar formation.
Silva et al. also explored the potential application of bioreactive nanoscaffolding in the treatment of central nerve regeneration. Their in vitro study reports on the encapsulating murine neural progenitor cells (NPCs) into a network of nanofibers produced by the self-assembly of peptide amphiphilic molecules (Fig. 3). The scaffolds were customized through adjustment in peptide sequences to have an outer surface that resembled the basement membrane protein laminin, which is known to increase neurite differentiation and growth. The challenge in neural tissue engineering is to induce NPCs to differentiate and grow as mature nerve cells. This is complex, because NPCs can also differentiate into astrocytes, oligodendrocytes, and ependymal cells (glial lineage). Astrocyte growth is undesirable because it is central to scar tissue formation.

In the study, cells in the nanofiber network resulted in 30–50% NPC promotion of cell differentiation into neurons, whereas only 1–5% of NPC differentiated into astrocytes (days 1 and 7). In contrast, control groups without nanoscaffolding had 5–15% neuron growth and 15% astrocyte differentiation. This study was the first to describe a nanoengineered material that resulted in a preferential differentiation of a desired cell type. In other has also been shown to be a positive cue in nerve stem cell neurite outgrowth and functional synapse formation, both of which are critical aspects to neurorepair.

The use of nanofiber scaffolding has also been investigated in the field of wound healing and skin regeneration. Rho et al. used electrospun type I collagen nanofibers to create biomimetic scaffolding for the attachment and spread of human keratinocytes. In the in vivo component of the study, a 1 × 1 cm full-thickness wound was surgically created in 12 rats. Six rats had a gauze pad applied with the collagen nanofibers, and six rats functioned as the control with untreated cotton gauze applied to the wound. Although late-stage wound healing were similar between control and experimental groups, microscopic examination revealed faster early-stage healing in the collagen nanofiber group. At 1 wk, tissue debris in the nanofiber group had been cleared, and capillary and fibroblast growth was prominent compared with in the control group, in which the wound surface was covered with fibrous debris overlying a dense infiltration of polymorphonuclear leukocytes. Yamaguchi et al. also showed that nanostructures stimulate keratinocyte growth in a skin injury model and increase inflammatory cytokine response to injury.

Given the extensive orthopedic interest in bone regeneration, much attention has been given to developing a scaffold that mirrors the hierarchical organization of bone. Bone is a composition of collagen fibers aligned in triple helices, upon which are layered hydroxyapatite crystals arranged along their c axes. Hartgerink et al. have developed a self-assembling fibrous scaffold of peptide–amphiphile molecules that resembles bone’s organic template. This scaffolding was designed to have a repetitive organization of phosphate groups and a surface-associated protein, fibronectin, to mimic bone’s ultrastructure. When pH is lowered to below 4, it self-assembles and disassembles when the pH is brought back to neutral. Surprisingly, the scaffolding directed the mineralization of hydroxyapatite crystals in the c-axis along the long axis of the fibers, mimicking the organization of natural bone. Nanocomposites resembling bone have been formed through other varied materials and mechanisms. Investigators have developed hydroxyapatite/collagen self-organizing matrices, hydroxyapatite/gelatin preparations, nanophase ceramics, and even carbon fiber surfaces. Importantly, these systems have all been tested for biocompatibility and have demonstrated good osteoconductivity in vitro.

Similar nanofibrous scaffolds have been used for biocompatible and biodegradable matrices for cartilage and muscle tissue engineering. Fur-
thermore, a nanofibrous scaffold has been used to promote selective differentiation in mesenchymal cells, with cellular viability, attachment, and proliferation of either adipose, cartilage, or bone, depending on the appropriate inductive agents. In contrast to the in vitro process, in vivo tissue-engineering techniques exploit the native environment to induce regeneration. Some tissues such as cornea (possessing a natural collagen scaffold) and bone marrow do not require an artificial scaffold. However, for the majority of tissues, especially when there is a large defect, scaffolding with or without cell seeding is necessary. An important consideration for future clinical applications is that self-assembling scaffolds have the distinct advantage of being injectable, with studies already demonstrating their assembly, safety, and tolerance in animal models.14,15,19

Therapeutic application of regenerative modalities may become commonplace in the future for physiatrists working with stroke, spinal cord injury, arthritic, or myopathic patients. Physiatrists, for example, may someday inject a partial anterior cruciate ligament tear, atrophied thenar muscle, or degenerated intervertebral disc with a regenerative combination of self-assembling scaffolding and pluripotent cells, and then monitor the complex interaction between regrowth of these tissues and improvement in clinical function.

In summary, varied tissue nanobioengineering models, combined with advances in stem cell research and an increased understanding of cytokine and protein factors influencing cell growth, are likely to translate into useful clinical tools to promote regeneration and successful transplantation in the foreseeable future.

Drug and Gene Delivery

Nano-based systems are in various stages of development for the delivery of genes and therapeutic agents. In conventional drug delivery, formulations based on simple physical properties (e.g., solubility) often control drug release. In contrast, nanosystems are more versatile synthetic chemical formulations that allow tailoring of particle size, with features that mimic biological structure. Nano-based systems offer several important advantages over conventional therapy. Drugs or genes can be protected in the vascular system and have controlled release at a chosen site of action. The targeted delivery can be achieved through interaction with cell receptors, transcription factors, or signaling proteins. Gene delivery and antigen display can be realized without inducing an immune response and with greater accuracy with respect to the target organ. Finally, the nano size of the devices allows for penetration of the blood–brain barrier (BBB) and stomach epithelium, and evasion of capture by the reticuloendothelial system.

Individualized nanosystems are based on several modifiable structures, including nanoparticles, nanoshells, and dendrimers. The advantage of using these structures is that they are inert, nontoxic, and biocompatible. Their surface and core are highly modifiable through conjugation with cellular epitopes and antibodies that bind to targeted cell membranes. For example, Thomas et al.29 report on one dendrimer system that was engineered with three distinct components: a folic acid as the targeting molecule, methotrexate as the chemotherapeutic drug, and fluorescein as the contrast probe. This dendrimer system permitted enhanced receptor-mediated exchange of the drug to the intended tumor cells, and tracking of its absorption and elimination.

Nanostructure surfaces can be engineered to be cationic or anionic, depending on the desired outcome of vascular or cellular penetration. The structures are also highly stable and can be used to protect easily denatured and less stable drug molecules. This property potentially allows for an intravenous, intramuscular, or subcutaneous drug to be packaged into an oral form.30 Finally, nanostructures can be magnetized or bound to a radiopaque dye so that drug release can be visualized through magnetic resonance imaging or computed tomography.31

Polymer therapeutics and nanocarriers show great promise in many aspects of medicine, with several formulations either approved by the U.S. Food and Drug Administration (FDA) or in clinical trials.32,33 These improvements, in combination with a greater understanding of genomics, cell transport systems, and molecular basis of pathology, will result in safer and more direct drugs to treat disease. Targeted drug delivery has been studied most widely for use in chemotherapy agents, with the potential for more specific action and decreased systemic toxicity. Proof of principal has been demonstrated in animal models of brain tumors,34 osteosarcomas,34 and esophageal cancer.35 In January 2005, a nanoformulation of the conventional chemotherapeutic agent paclitaxel was FDA approved. The drug (Abraxane) is approved for use in patients with metastatic breast cancer. By virtue of its molecular size, the drug binds to naturally occurring albumin, resulting in better tolerance and fewer side effects. In fact, steroids no longer need to be administered before and after administration, as is needed with conventional paclitaxel.

Using the established principles in nanostructured drug delivery, the field of physical medicine will benefit from the general application of these engineered drugs for the systemic treatment of our patients. However, there also is potential for a more direct application of nanostructured drugs in
particles loaded with ibuprofen also have been ideally suited to treating chronic conditions. Nanocacy, a nanoformulation of glucocorticoid may be cellular site of action determine its therapeutic effi-
cy. The key attribute of these treelike structures is a synthetic system.45 There are several barriers to achieving successful gene delivery, including transport of the nucleic acid in the circulation, entry into the cell, dissociation from the vector, transit from cytoplasm to nucleus, entry into nucleus, and expression of the transgene.46 Liposomes, linear polymers, microbubbles, and nanoparticles have been used as platforms for gene delivery, with dendrimers studied most extensively. In contrast to gene delivery through use of viral vectors, these agents can be safely administered without eliciting an immune response in hosts.

Dendrimers are very adaptable, allowing for a broad range of molecules with varying functionality. The key attribute of these treelike structures is the density of terminal groups (or branches), allowing for a high surface ratio that offers multiple attachment sites for conjugation of the gene, targeting segment, or detection agent. Much of the current research in dendrimer gene delivery involves the targeting of solid tumors, with successful in vivo delivery and transfection demonstrated in several studies.46,47 In addition to simple nanostructures, more complex multicomponent systems are being constructed. These various forms of “artificial viruses” have a layered assembly of a cationic core composed of functional peptides (to aid in nuclear transport), surrounded by an anionic shell of flexible polymers and target ligands (which can induce receptor-mediated endocytosis). Although the “artificial virus” is in the initial phases of discovery, these elaborately engineered systems are being constructed and tested by multiple laboratories worldwide.

For physiatrists, these nanotechnologies may impact the treatment of patients with genetically

rehabilitation medicine. For example, one of the most commonly used drugs in musculoskeletal care is the localized injection of glucocorticoids, which have intracellular sites of action. Cytoplasmic concentrations of steroids are maintained only transiently when used in solution. Hence, the ther-
apeutic effect is seen only for a short duration. Panyam and Labhasetwar36 developed a nanoparticle formulation of dexamethasone and tested its intracellular delivery and effectiveness in inhibiting cell proliferation compared with the effectiveness of standard solutions. In cultures of human vascular smooth muscle, the nanoparticle formulation maintained sustained intracellular drug levels, whereas the levels dropped rapidly in the cells that were treated with the drug in solution. Of the two nanoparticle formulations used in the study, the higher-dose formulation demonstrated greater, more sustained intracellular drug levels (35-fold greater intracellular concentration on day 8 compared with the lower dose). Also, the intracellular drug levels for the two formulations correlated with their antiproliferative activity. Significantly higher, more sustained (up to 12 days) inhibition of cell proliferation was obtained with dexamethasone-loaded nanoparticles compared with the standard solution (up to 5 days). Electromyography also showed that nanoparticles were detected in the cytoplasm 14 days after incubation. Importantly, the nanoformulation of dexamethasone accumulated in greater concentrations intracellularly and had a slower and more effective sustained release than the conventional solution. Because the dose and duration of a drug’s availability at the intracel-

ular site of action determine its therapeutic effi-
cacy, a nanoformulation of glucocorticoid may be ideally suited to treating chronic conditions. Nanoparticles loaded with ibuprofen also have been successfully fabricated, with further study required to determine their efficacy compared with that of standard formulations. Alternative forms of deep heat are achievable through new photothermal composites using nanoshells. These optically active nanoparticles consist of a conductive metal layer (usually gold) surrounding a dielectric core and can be altered to a desired shell thickness and particle size.38 On absorbing a near-infrared light at a tunable resonance, the nanoshells emit a dose of heat to a defined tissue volume. Feasibility studies for this technique for thermal tumor ablation have been conducted in vitro and in vivo, in murine colon carcinoma and human breast epithelial carcinoma models.

Another application of nanostructured drugs in rehabilitation medicine involves chronic degenerative neurologic diseases. In treatment models of neurodegenerative disorders such as Alzheimer disease and amyotrophic lateral sclerosis, fullerenes have been shown to have powerful neuroprotective antioxidant qualities. As free radicals are collected by fullerenes, there is a reduction in excitotoxicity and exocytosis occurring with neuronal injury. Promising animal studies in Parkinson disease models have indicated improved motor performance, increased survival, and preserved dopamine production.42 It has also been suggested that free radicals contribute to multiple cellular injury cascades in stroke, ataxia telangiectasia, traumatic brain injury, and spinal cord injury, representing a variety of potential areas of fullerene-based treatment research.

Gene therapy is another developing field with potential applications in rehabilitation medicine. For accurate gene delivery, several nanostructures are being investigated as targeting systems to overcome the limitations and side effects of viral vectors. To introduce a nucleic acid into a cell, the challenge is to design a system that incorporates the transfection efficiency of a virus with the ease of manufacture, safety, and targeting capability of a synthetic system.45 There are several barriers to achieving successful gene delivery, including transport of the nucleic acid in the circulation, entry into the cell, dissociation from the vector, transit from cytoplasm to nucleus, entry into nucleus, and expression of the transgene.46 Lipo-

omes, linear polymers, microbubbles, and nanoparticles have been used as platforms for gene delivery, with dendrimers studied most extensively. In contrast to gene delivery through use of viral vectors, these agents can be safely administered without eliciting an immune response in hosts.

Dendrimers are very adaptable, allowing for a broad range of molecules with varying functionality. The key attribute of these treelike structures is the density of terminal groups (or branches), allowing for a high surface ratio that offers multiple attachment sites for conjugation of the gene, targeting segment, or detection agent. Much of the current research in dendrimer gene delivery involves the targeting of solid tumors, with successful in vivo delivery and transfection demonstrated in several studies.46,47 In addition to simple nanostructures, more complex multicomponent systems are being constructed. These various forms of “artificial viruses” have a layered assembly of a cationic core composed of functional peptides (to aid in nuclear transport), surrounded by an anionic shell of flexible polymers and target ligands (which can induce receptor-mediated endocytosis). Although the “artificial virus” is in the initial phases of discovery, these elaborately engineered systems are being constructed and tested by multiple laboratories worldwide.

For physiatrists, these nanotechnologies may impact the treatment of patients with genetically
inherited disorders (i.e., muscular dystrophies, spinal muscular atrophy, Huntington disease) and disorders of cellular regulation (i.e., osteosarcoma, astrocytoma).

Drug transport across the BBB is also a major area of nanodrug delivery research and may be of particular interest to physiatrists working with central nervous system disorders. The tight junctions of the BBB strictly limit exchange to hydrophilic compounds, small molecules, and charged molecules. A major challenge for the treatment of most central nervous system disorders is overcoming the difficulty of delivering therapeutic agents. The advantage of a nanodrug formulation is that a nanoparticle, liposome, or dendrimer can mask BBB limiting qualities (such as surface polarity) while stabilizing the drug for transport. The small size allows a nanodrug to also be endocytosed by endothelial cells and, if functionalized, take advantage of receptor-mediated endocytosis, an effective alternative to passive diffusion. In certain cases, peripheral toxicity also can be reduced. Drug packaging can protect a drug from rapid enzymatic degradation and elimination in the brain, allowing for sustained release. Some rehabilitation-relevant examples of central nervous system drug delivery using nanoparticles include amitriptyline, tubocurare, and valproic acid. A full discussion of this developing area of pharmaceutical research is outside the scope of this paper, and the reader is referred to several comprehensive review articles of nanotechnology-based drug delivery across the BBB for a more detailed analysis.

**Imaging**

Advances in nanotechnology have given rise to unique possibilities for studying disease in a non-invasive manner. Established imaging techniques such as magnetic resonance, radionucleotide (i.e., positron emission tomography), and optical imaging have improved steadily in the past two decades. State-of-the-art nano-based imaging systems have evolved from these established techniques to provide even greater detail and amplification for in vivo imaging of specific target tissues, gene activity, and molecular processes. The technologies are based largely on the use of quantum dots (QDs) and nanoparticles, which can be conjugated with antibodies and recombinant proteins to target DNA, cell-surface receptors, and intracellular enzymes. Because of their small size, they are also more efficient than conventional biological stains, contrast agents, or fluorescent probes in traversing cell membranes and diffusing through interstitial spaces.

Semiconductor nanocrystals, also referred to as QDs, have been well studied for their potential role in biological imaging. QDs have tunable emission from ultraviolet to infrared. Depending on particle size (2–8 nm), they will always emit the same “color” frequency (Fig. 4). This property makes...

---

**FIGURE 4** Sensitivity and multicolor capability of quantum dot (QD) imaging in live animals. The right-hand images show QD-tagged cancer cells emitting green, yellow, or red light. Approximately 1–2 million beads in each color were injected subcutaneously at three adjacent locations on a host animal. Reprinted with permission from Gao X, Cui Y, Levenson RM, Chung LW, Nie S: In vivo cancer targeting and imaging with semiconductor quantum dots. Nat Biotechnol 2004;22:969–76 (www.nature.com).
QDs especially useful as biomolecular probes for use in imaging and diagnostic assays. For QDs to be used in the study of biological processes, they must be altered to optimize water solubility. Several strategies have been employed, including conjugation with water-soluble ligands, silanization, and encapsulation within micelles.

Unlike other molecules that fluoresce when excited at certain wavelengths, QDs have a broad excitation spectrum, which makes it possible to excite many QD “colors” using a single laser wavelength. This enables simultaneous probing for multiple biosensing markers. QDs also show less photobleaching than conventional probes, higher luminescence, and longer absorption with minimal degradation in emission. QDs are so bright that it is possible to detect a cell carrying only one crystal. The extremely high photostability of QDs allows for real-time monitoring or tracking of intracellular processes over long periods of time. Furthermore, QDs have a high surface area-volume ratio, allowing for complex nanosystem design and functionalization. To target specific tissues, QDs have been bound to peptides that recognize various cellular markers or intracellular processes. In previous studies, QDs have been bioconjugated into probes suitable for imaging and staining of prostate cancer cells, tumor blood vessels, lung tissue, and fibroblasts. QDs also can be used as fluorescent tracers for uptake studies of dynamic processes such as embryogenesis and angiography. These techniques have shown high in vivo selectivity and greater power and detail than conventional methods.

QDs have been studied most extensively in cancer cell imaging. Of interest to physiatrists, QDs also have been used as probes to study neuronal processing. For example, Dahan et al. used QDs to selectively track glycine-receptor diffusion in spinal cord neurons (glycine receptors are the primary inhibitory neurotransmitter receptor in the spinal cord). The significance of using a QD as a probe is the ultrasensitive measurement it allows. More powerful than larger beads or small fluorophores, the QDs used by Dahan et al. were able to accurately record the mobility of individual molecules at the neuronal surface. Dahan et al. also were able to acquire both fluorescence and electro-myographic images using the same probe, a powerful way to establish high-resolution localization. QDs also have been used to study neuronal differentiation response to nerve growth factor and mobility of AMPA receptors.

Nanoparticles, especially magnetic nanoparticles, have also been studied extensively for their potential use as improved contrast agents and biological probes. These ultrasmall contrast agents have unique qualities that allow for enhanced imaging of occult metastases. Two FDA-approved agents for detection of occult liver (ferumoxides) and lymph node (ferumoxtran-10) metastasis are currently in clinical use. When given intravenously, these ultrasmall particles of iron oxide (USPIO) slowly extravasate into the interstitial space, are internalized by macrophages, and are taken up into the reticuloendothelial system of lymph nodes. On T2-weighted magnetic resonance image sequences, uptake of these results in a loss of signal intensity (darkening). Metastatic nodes block uptake of USPIO, and the difference can easily be detected visually. Magnetic nanoparticles have been demonstrated to function as magnetic resonance enhancers for lymph node detection of prostate cancer, endometrial cancer, cervical cancer, abdominal cancer, and brain tumors. They also can be delivered into cells to enhance magnetic resonance imaging of individual cell movement, gene expression, and progenitor cell distribution and differentiation. Better targeting to specific tissues is also possible through attachment to dendrimers or encapsulation within liposomes.

Nanoparticulate contrasts have the unique advantage of enhancing the magnetic resonance signal from the tissue and cell level compared with the conventional blood-pool agents that enhance signal from the vascular system. The potential for using iron oxide–based contrasts for non–cancer-related imaging is being investigated. For example, ferumoxtran-10 (Combidx, Advanced Magnetics, Inc.) has been studied for targeting central nervous system inflammatory lesions. Twenty-three patients with demyelinating lesions (i.e., multiple sclerosis, acute disseminated encephalomyelitis), vascular lesions (i.e., stroke), and hematopoietic neoplasms (i.e., lymphoma) received ferumoxtran-10 magnetic resonance scans in addition to standard magnetic resonance imaging as part of a diagnostic workup. Whereas demyelinating lesions failed to show superior enhancement, subtle stroke lesions showed greater signal intensity on T2-weighted images using ferumoxtran-10 compared with gadolinium. Also, one symptomatic cavernous venous vascular malformation did not enhance with gadolinium, but it did enhance intensely with ferumoxtran-10. Hematopoietic neoplasms all enhanced with the USPIO, with varied results in signal intensity with each patient compared with gadolinium. USPIOs also have been used for imaging inflammation in animal models of spinal cord injury, multiple sclerosis, and oligodendrocyte remyelination, successfully demonstrating in vivo USPIO labeling of cellular infiltration.

For physiatrists, the general application of these new imaging modalities for diagnosis and therapeutic interventions may prove useful. In addition to the active research previously mentioned, potential applications include the use of nano-
based imaging to study (a) protein and lipid changes in the brain after a stroke or traumatic brain injury, which could then be correlated functionally; (b) nociceptive and inflammatory factors in spine and musculoskeletal disorders to accurately localize sites of pain generation and inflammation; or (c) nerve root, axon, or myelin components to assess degeneration or recovery as an adjunct to electrodiagnostics. New imaging techniques with nanomaterials represent an advance that may increase accuracy in visualizing pathology and inflammation, labeling large numbers of cells, and targeting specific tissues and dynamic cellular processes.

**Diagnostic Assays**

Current diagnostic techniques such as Western blots, enzyme-linked immunosorbent assay, cell culture, latex agglutination, and polymerase chain reaction often require purification, separation, and incubation steps that require increased time, effort, and cost for processing. Development of miniaturized devices using nanotechnology may provide faster, more accurate, and, ultimately, less expensive technologies for diagnosis of disease and microbial pathogens. The main goals of current nano research are to (a) create high-throughput assays that are rapid and sensitive and that show real-time processes; (b) develop diagnostic tools that do not require additional processes such as target amplification or use of radioactive labels; and (c) make assays broadly available for use at the bedside, in contrast to current methodologies (e.g., culture, polymerase chain reaction) that require dedicated technicians in centralized hospital or commercial laboratories.

Currently, a diverse array of nanotechnologies have been adapted for potential application as commercial diagnostic assays. A gold nanoparticle sandwich assay has been well described in the literature for the ultrasensitive detection of proteins, genes, and nucleic acids (Fig. 5). First reported in 1996, the technique involves oligonucleotide attachment to two sets of nanoparticles (magnetic and gold) that sandwich the target. Then, the target compound is isolated with magnetic separation and identified. Nanosphere Inc. (Northbrook, IL) has developed an automated platform based on this technology (Verigene System) and has established FDA approval for commercial clinical testing of nucleic acids. Bio-bar-code is another sandwich assay supported by the same company and has been studied for the detection of proteins and target DNA. It has been shown to be as sensitive as conventional polymerase chain reaction for DNA detection (even differentiating single base mismatches), but it does not require the standard enzyme-amplification steps. The Clear-Read system has been developed to target single nucleotide polymorphisms and can detect DNA sequences in a single drop of blood or buccal swab. Several other nano-based methods have been used for detection of genes and SNPs, including fluorescent QD assays, engineered nanopores, and dip-pen nanolithography for creation of high density gene-chips.

To analyze non-DNA biomolecules and proteins, a wide diversity of diagnostic techniques...
has been described using nanoparticles, QDs, and nanoshells. Georganopoulou et al.\(^93\) describe a nanoparticle-based method for detection of the Alzheimer disease biomarker amyloid-\(\beta\)-derived diffusible ligands. By using a nanoparticle assay, the investigators were able to quantify the antigen in the cerebral spinal fluid of test subjects at a concentration beyond the sensitivity of conventional techniques. Wang and colleagues\(^94\) recently reported their use of silicon nanowire field-effect transistors to identify and quantify molecular inhibitors of adenosine triphosphate or substrate proteins to tyrosine kinase. Similar to calorimetry and surface plasmon resonance, the technology of silicon nanowire field-effect transistors is a label-free way to detect binding and unbinding of proteins to their ligands. The advantage of silicon nanowire field-effect transistors is that it is more sensitive than conventional methods and requires a far smaller sample.\(^95\)

The highly luminescent properties of QDs also make them ideal for large-scale labeling of molecules and rapid target identification. Several groups have used QDs as biological probes.\(^96\) In fact, five differently sized QDs with tunable emission could yield 10,000 codes detectable through spectroscopic measurement. Each QD can be its own “chemical lab,” detecting and analyzing a unique compound and resulting in faster and more flexible high-throughput assays.\(^97\) Nanoshell probes share similar characteristics to QDs and feature high biocompatibility and near-infrared emission. Nanoshells are easily conjugated to antibodies and can be fabricated to desired optical resonances. This technique shows promise for the development of a simple, in situ whole-blood immunoassay comparable with conventional enzyme-linked immunosorbent assay and latex agglutination.\(^98\)

Nanofluid techniques have also improved conventional sampling methods. This has been especially true for improvement in mass spectroscopy and liquid-phase separation techniques. One critical step of mass spectroscopy is the liquid junction interface, which involves applying a voltage to the sample as it is injected via an electrospray into the inlet of the spectrometer. Many different designs have been formulated for the electrospray tip, which requires a precise jet and alignment. Gold-coated electronanosprayers are widely available and have been shown to result in more sensitive and rapid interface of the compound.\(^99\) New commercial Lab-on-a-Chip technology takes advantage of nanoflow improvements to create smaller, more effective diagnostic tools that require minute samples to run multiple analyses.\(^100\) Fluid techniques also can isolate individual molecule nanoconcentrations. The ability to analyze precise samples has allowed for the detection and quantification of high-concentration enzymatic processes such as synthesis of double-stranded DNA by DNA polymerase.\(^101\) Furthermore, advancements in the nanowiring of electronic circuits allow for smaller, more easily implanted analytic tools.\(^95\)

For physiatrists, these new systems of DNA, enzyme, protein, and antibody analysis may significantly impact the efficiency and cost of diagnosing disease. Clinical testing of cytogenetic disorders (e.g., chromosomal variants associated with Charcot–Marie–Tooth Disease), enzyme abnormalities (e.g., elevated creatine phosphokinase in Duchenne/Becker muscular dystrophy), protein over/underproduction (e.g., overproduction of M-protein in monoclonal gammopathy), or disease-related antibodies (e.g., rheumatoid factor in rheumatoid arthritis) are all possible through currently available nanotechnology-based assays.\(^102,103\) In the inpatient setting, nano-inspired tools will allow the physiatrist to gain rapid and accurate data on microbial pathogens causing acute or chronic infections, such as pneumonia and urinary tract infections.\(^104\) Proof of principle with different nano-based diagnostics has already been established for bacteria (E. coli,\(^105\) S. aureus\(^87,106\)) and viruses (influenza, adenovirus,\(^107\) herpes simplex\(^108\)). Furthermore, nanotechnology may allow for the early detection of emerging antibiotic resistance.\(^87\) Importantly, these diagnostic techniques overlap with those related to drug delivery and imaging and demonstrate the versatility of nanocompounds in their application to the biological sciences and clinical medicine.

**Texturing/Biointerfaces**

Nanotexturing and nanocoating are manufacturing techniques applied to inert materials to establish bioactivity and to enhance biocompatibility. The insertion of nonorganic implantable devices may have limited efficacy and functional life because of the body’s intrinsic responses to any object recognized as “foreign.” Problems of tissue ingrowth, inflammation, infection, and thrombogenesis need to be addressed to optimize biocompatibility.

Nanotexturing is a method to enhance surface topography to include nanoscale features such as pits, islands, or ribbons of various depths. There are several methods to produce nanotopography on various polymer, metal, and ceramic surfaces, including electron beam nanolithography, self-assembling systems, microcontact printing, and injection molding.\(^109\) These processes embed a biological matrix to an otherwise nonbiological surface. There is growing evidence that nano features strongly affect surrounding cell growth, proliferation, and differentiation. Human fibroblasts, for example, were shown to increase cell spread, formation of filopodia, and differentiation when grown on nanoscale island cytoskeleton compared
with fibroblasts grown on a flat surface.\textsuperscript{110} Osteoblasts also have been shown in multiple studies to increase cell adhesion and growth on surfaces that resemble the nanostructured geometry of hydroxyapatite,\textsuperscript{22,111–113} an important consideration for the osseointegration of orthopedic implants.

Nanocoating is the addition of a bioactive component to a surface, changing the surface chemistry of the implant to be either more cell attractant or cell resistant. The challenge to researchers is to develop a physiologically stable biointerface that can be modulated to the desired function. For example, the ideal nanocoating for an orthopedic implant would be one that would be resistant to oxidative, hydrolytic, and mechanical forces, yet able to attract osteoblasts directly to the surface of the metal to improve osseointegration. Methods of nanocoating include self-assembly, layer-by-layer assembly, and electrolytic deposition. Various biomaterials have been used in the nanocoating process, including collagen,\textsuperscript{114} collagen composites,\textsuperscript{115,116} proteins,\textsuperscript{117} and inorganic polymers.\textsuperscript{118} As in texturing, the result is a surface that can increase desired cell adhesion and growth for implants.\textsuperscript{115} A surface can reduce unwanted biological activities for prosthetic devices in vivo, such as scar tissue\textsuperscript{113} or thrombus formation.\textsuperscript{118}

In addition to orthopedic implants, the areas that most directly impact rehabilitation medicine are the nanocoating of neural implants and cardiovascular stents. Ai et al.\textsuperscript{119} describe a process of applying a multilayered nanofilm composed of various biocompatible elements, including laminin and fibronectin, onto silicone rubber tubes. He et al.\textsuperscript{120} also report on a technique of layered assembly of neurointegrative coating for silicon arrays. Currently, silicone tubes are the implant interface for ventricular drainage shunts, spinal cord electrical stimulators, cochlear implants, and peripheral nerve-reconstruction scaffolds. Importantly, in both studies above, neurons demonstrated cell adhesion, growth, and morphological differentiation comparable with that from standard, thicker coatings, without cytotoxicity. These nanocoatings allow for biocompatibility without changing the size characteristics of the device. This may be critical in the use of neural implants as prosthetic devices for chronic disease (such as cerebral palsy), where drug delivery or sensory stimulation will require long-term integration of the device with surrounding neural tissue. In addition, these neurointegrative coatings may impact those physiatrists working in the areas of direct brain interfacing for control of prosthetics and the development of biohybrid limbs. For cardiovascular stents, nanocoating has been shown to enhance plaque prevention\textsuperscript{120} and drug delivery\textsuperscript{121} in animal models of carotid and coronary artery injury. Arterial stents are widely used in the clinical management of stroke and coronary artery disease, with nanocoatings representing a more advanced and versatile system than current drug-eluting technology.

The potential impact of nanotexturing and nanocoating to the field of physiatry is significant; given the diversity of implantable devices used in rehabilitation-related fields (orthopedic, neurologic, cardiovascular, etc.). In summary, the goal of the evolving technologies of nanotexturing and nanocoating is to achieve optimal integration of living and synthetic components to restore normal tissue and cell function for skin, vessel, nerve, bone, ligament, and tendon. Wherever an inorganic/organic interaction exists, nanotechnology will be studied to control this interface.

CAUTIONARY EPILOGUE

Nanotechnology is developing rapidly, with the applications of the field being examined by an increasing number of scientists, venture capitalists, government agencies, and politicians. As the attention has increased, so has the debate. Is nanotechnology the industrial revolution of the 21st century? Or have the potential benefits been oversold by nanotechnology enthusiasts? It remains to be seen whether the initial excitement will result in meaningful clinical applications and returns on financial investment.

As with all new biotechnology fields, serious concerns have been raised regarding the environmental hazards and health risks regarding nanostructures. The uncertainty is based on the very aspect of nanotechnology that makes it so appealing: the potential outcomes are unknown and nearly impossible to predict. Concern for wildlife was publicized in a study of potential water-borne effects of buckyballs in which it was demonstrated that the lipophilic uncoated fullerenes selectively translocate into the brain of large-mouth bass and result in oxidative damage.\textsuperscript{122} In humans, nanoparticles can be rapidly transported into the bloodstream through the lungs, intestines, and even the skin.\textsuperscript{123} Inhaled nanoparticles have been associated with inhibited cellular signaling\textsuperscript{124} and oxidative stress in the liver.\textsuperscript{123} Nanotubes also have been reported to induce granuloma formation in the lungs of rats and mice.\textsuperscript{125,126}

Because the complete toxicity of nanostructures to humans and the environment are unknown at the present time, many argue for rigorous safety testing before the initiation of widespread use of nanomaterials in medicine and industry. This issue has received federal funding, with several grants distributed by the U.S. Environmental Protection Agency and the National Toxicology Program to investigate the effects of nanomaterials on aquatic systems, air pollution, and exposure limits for workers.\textsuperscript{127} The

March 2007

Introduction to Nanotechnology
National Nanotechnology Initiative alone has budgeted $106 million to fund the study of the ethical, environmental, and health implications of nanoparticles. Others have argued for government standardization of assays and characterization of materials to protect the public and to facilitate quality control of manufacturing techniques and reproducible research. Currently, there are no specific regulations pertaining to nanomaterials. This reality heightens the fears of those concerned about the unrestrained and uncontrolled growth of the field. The subject of moratoriums on research has been introduced at international conferences to encourage responsible development, but, as yet, nothing has curtailed the private and academic pursuit of this emerging field.

Increasingly, private and government-funded funds are being poured into the development of novel nanoscale products. However, there are several obstacles to overcome before theoretical ideas can be translated into biomedical reality. Many products may be complementary to current microtechnology materials, but they may not have improved performance. Manufacturing costs may be prohibitive and may not warrant the investment needed for production. In essence, venture capitalists are avidly courting and supporting many new nanotechnology start-up companies, but the usefulness and affordability of many nanoproducts has not yet been determined.

One aspect of the field that cannot be argued is that nanotechnology has captured the imagination of scientists, physicians, and the public. As the field continues to grow and impact more areas of medicine, it is important for those in the field of physical medicine and rehabilitation to be familiar with the potential biomedical benefits and risks of nanoscale products. Judicious research may elucidate those aspects of nanotechnology that improve the lives of patients with disabilities and translate to new methods of clinical practice.

ACKNOWLEDGMENTS

We would like to thank Dr. Nadrian Seeman (New York University, Department of Chemistry, DNA Nanotechnology Laboratories) and Jeffrey Schwartz (Princeton University, Department of Chemistry) for kindly reviewing the manuscript and providing valuable feedback.

REFERENCES

25. Li WJ, Tuli R, Okafor C, et al: A three-dimensional nano-


100. Agilent Technologies I. Available at: http://www.chem.agilent.com


122. Ober dorster E. Manufactured nanomaterials (fullerenes, C60) induce oxidative stress in the brain of juvenile largemouth bass. *Environ Health Perspect* 2004;112:1058–62


129. Fear and loathing; some of the worries about nanotechnology are rational, some not. *The Economist* 2005;374:7–9
Spinal Cord Injury Associated with Thoracic Osteoporotic Fracture

ABSTRACT


This report details a case of sudden neurologic deficit attributable to acute thoracic fractures associated with senile osteoporosis. A 73-yr-old female patient with a history of occasional back pain during the past 4 mos had sudden thoracic vertebral fracture with spinal cord injury. The patient, who had a benign past medical history, had not been evaluated for osteoporosis.

Thoracic spine radiographs showed a compression fracture at T8. Thoracic magnetic resonance imaging exposed a compression fracture at T7–T8. She was treated operatively. She was found to have spinal cord injury with American Spinal Injury Association classification C (T7), and she had poor sitting balance. She was discharged in a wheelchair and was administered clean intermittent catheterization every 6 hrs. Six months after discharge, she ambulated with a walker and had spontaneous micturition. Vertebral fractures are a common presentation of senile osteoporosis. The risk of neurologic impairment attributable to vertebral fracture is a rare but potentially severe complication. Besides medical therapy and suitable rehabilitation programs, surgical treatment is an integral part of the management of patients with osteoporotic vertebral fractures.

Key Words: Osteoporosis, Vertebral Fractures, Spinal Cord Injury

The osteoporotic vertebral fractures that occur spontaneously or after minimal trauma in the course of daily living can be a major cause of morbidity in elderly people. Twenty percent of women older than age 65 have had one or more vertebral fractures.1 In many of these cases, the incidence of vertebral fracture increased with rising age in both sexes. Most of these fractures are subclinical, and the majority are thought to heal without complications.2 Spinal fractures can adversely affect quality of life by causing pain, reducing physical function and mobility, and affecting activities of daily living.2 Despite severe spinal deformities, nerve root– and spinal cord compression are rare.3–6 Neurologic impairment generally occurs between 1 mo and 1.5 yrs after fracture. Sudden-onset spinal cord injury is extremely rare.7 We present a case of...
immediate spinal cord injury attributable to thoracic fractures associated with senile osteoporosis.

**CASE REPORT**

A 73-yr-old female patient with a history of occasional back pain during the past 4 mos was admitted to a neurosurgery clinic because of sudden weakness in the lower extremities and urinary incontinence. The patient resided with her family, and they denied any history of recent trauma. The patient had not attended an outpatient clinic and, therefore, had not been evaluated for osteoporosis. The patient had a benign past medical history, with the exception of diabetes mellitus and hypertension. Erythrocyte-sedimentation rate, complete blood count, serum levels of calcium and phosphate, alkaline phosphatase, protein electrophoresis, hepatic enzymes, tumor markers, and abdomen ultrasonography were all normal. Direct radiographic evaluation revealed a compression fracture at T8 (Fig. 1). Magnetic resonance imaging illustrated a compression fracture at T7–T8 and cord compression (Fig. 2A). Computed tomographic imaging showed bony fragments in the spinal canal and spinal cord compression (Fig. 2B). After corpectomy, decompression, allografting, and injection of methacrylate at T7–T8, instrumentation was performed from T6 to T9 (Fig. 3). Postoperatively, the motor and sensory deficits recovered partially. Subsequent histology showed no evidence of neoplasia or infection and was consistent with early new-bone formation. Intraoperative cultures also were negative.

The patient was admitted to our clinic postoperatively with the diagnosis of nontraumatic paraplegia. She was found to have a poor sitting balance, and there was a stage II pressure sore on the sacrum. She also complained of fecal and urinary incontinence. Muscle tonus was mildly increased and the deep tendon reflexes were hyperactive. The upper-extremity motor scores were 25, bilaterally. The lower-extremity motor score was 12 on the right and 0 on the left according to the American Spinal Injury Association (ASIA) classification. The light-touch scores were 40 bilaterally, and pinprick scores were 42 bilaterally. The patient was evaluated as ASIA C (T7). The motor FIM score was 23 and the cognitive FIM score was 35. The pressure sore on the patient's sacrum was covered with a hydrocolloid wound dressing. She was mobilized with a thoracic orthosis. She participated in a multidisciplinary rehabilitation program. A tilting program was started. She was given range-of-motion exercises, posture exercises, balance and coordination training, progressive resistive exercises, and active assistive exercises. As exercising precipitated the patient's hypertension, her antihypertensive therapy was adjusted. When sitting balance was accomplished, assisted standing was achieved at the parallel bars. Standing balance was practiced with a posterior shell and an assisting dorsiflexor. Duration of the rehabilitation stay was 2.5 mos. Unfortunately, the patient could not achieve any type of ambulation because of very weak trunk muscles and upper extremities. The patient's hypertension was also a negative factor. Urodynamic examination showed a maximum cystometric bladder capacity of 200 ml and a maximum detrusor pressure of 42 cm H2O with detrusor–sphincter dyssynergia. Oxybutynin hydrochloride was administered twice a day. Bone mineral density evaluation by dual-energy x-ray absorptiometry was performed. Osteoporosis and fracture risk was detected on lumbar (L1–L4: 0.589 g/cm², T score: −3.99, z score: −1.67) and hip (neck: 0.665 g/cm², T score: −3.42, z scores: 1.56). For osteoporosis, 70 mg/wk of alendronate sodium, 0.25 μg twice a day of alpha-calcidiol, and 500 mg of elemental calcium daily were administered. The lower-extremity motor score at discharge was 15 on the right and 10 on the left. She was still classified as ASIA impairment scale grade C. The motor FIM score at discharge was 35. The patient was discharged in a wheelchair, clearly in need of assistance during transfers and administration of clean intermittent catheterization every 6 hrs, with a bladder capacity of 250 ml.

Urodynamic examination, which was done at 6 mos after discharge, revealed that the patient's maximum cystometric bladder capacity increased from 200 to 475, and her maximum detrusor pressure decreased from 42 to 26 cm H2O. Six months after discharge, the lower-extremity motor score

---

**FIGURE 1** Lateral radiograph of the thoracic spine shows T8 compression fracture.
was 19 on the right and 18 on the left. She was now classified as ASIA impairment scale grade D, and she ambulated with a walker. The motor FIM score was 52. She had spontaneous micturition and defecation.

DISCUSSION

Back pain is a common complaint in the elderly age group. Vertebral osteoporosis, which is related to compression fractures and mechanical derangement of the spine, contributes to this pain. However, signs and symptoms of vertebral fractures may be mimicked by metabolic diseases such as hyperparathyroidism, hyperthyroidism, and multiple myeloma or other malignancies. Therefore, vertebral fractures necessitate comprehensive evaluation to rule out early etiology. Our patient had an acute paraplegic state after insidious back pain, with no history of trauma. The diagnosis of osteoporotic fracture was based on normal laboratory test results, magnetic resonance imaging, and postoperative histopathologic evaluation.

Compression fractures occur most often at the lower thoracic and upper lumbar spine, followed by the middle thoracic and lower lumbar spine, and, rarely, in the upper thoracic spine. The importance of angular deformation and the shape of the compression fracture are the other factors that predispose to neurologic complications. The loss of parallelism between the anterior and posterior walls of the vertebral body that, in traumatic fractures, implies a major risk of neurologic complications, is rarely seen in osteoporotic collapse. Moreover, when a wedge deformation occurs, it is usually located at the anterior part of the vertebral body. This probably explains the usual absence of neurologic compression. Kempinsky has reported two cases at the level of T5–T6 in which neurologic symptoms occurred a few weeks after the trauma (i.e., a fall on the buttocks). Subsequently, Salomon et al. have reported a case in which there was a T5 compression fracture with progressive neurologic impairment following trauma. Since then, there have been a few more reports in the literature, but none of these have shown immediate onset of paralysis. The only case in which there was acute neurologic impairment attributable to an osteoporotic fracture was reported by O’Connor et al. They reported a case of complete thoracic spinal cord injury with acute T5 fractures. The patient had continuous oral corticosteroids for more than 40 yrs for both rheumatoid arthritis and restrictive lung disease. The patient, who had not been evaluated and treated for osteoporosis, was managed nonoperatively because of her extremely osteo-

FIGURE 2 A, Sagittal T2-weighted fast spin-echo image shows compression fracture at T7–T8 and cord compression. B, Computed tomographic image shows displacement of bony fragment into the spinal canal with compression of the spinal cord.

FIGURE 3 Postoperative radiograph demonstrating stabilization extending from T6 to T9.
Spinal fracture in elderly patients is a typical manifestation of osteoporosis. Although fractures are generally associated with pain, it is important to consider the possibility of vertebral involvement in this population, particularly when there is an obvious progression of kyphotic deformity without any discomfort. Despite severe spinal deformities, neurologic injury is generally of a very low incidence. Besides appropriate medical therapy and a rehabilitation program, surgical treatment is an integral part of the management of patients with vertebral fractures.

REFERENCES
RE: RACIAL DISPARITIES IN ACCESS TO CARDIAC REHABILITATION

To the Editor: The article “Racial Disparities in Access to Cardiac Rehabilitation” by Gregory et al. demonstrated that black patients with heart disease were less likely to be referred for cardiac rehabilitation than white patients with similar clinical presentations. The authors termed this underuse of cardiac rehabilitation by black patients a “disparity.” This article highlights two important issues: the interpretation of data showing differential use of healthcare services, and the overall low rate of referral for cardiac rehabilitation despite its documented benefits.

The term “disparity” is frequently misused in the medical literature. Rathore et al. propose five formal criteria that should be fulfilled before defining a case of differential use in services by race as a disparity. These criteria include comparability between groups in eligibility criteria, clinical exclusion criteria, and preferences; lack of confounding (by factors such as demographics and social characteristics); and important consequences of not using the service in question (such as rehospitalization, morbidity, and mortality).

Gregory et al. provide data and discussion points that satisfy several of these suggested requirements. Black and white patients seem to be equally eligible for cardiac rehabilitation, with no systematic differences in exclusions. Literature in the field indicates that cardiac rehabilitation is effective. The authors adjusted for several potential confounders, although the attenuation of the odds ratio for race from 2.5 to 1.8 with adjustment indicates that the effect of race may not necessarily constitute an important difference, but not necessarily a disparity. Although Gregory et al. acknowledge the unavailability of data on patient preference as a limitation, they call the difference observed in their study a disparity. We should probe further about the basis of patients’ preferences and acknowledge that preferences themselves may reflect internalized oppression. It is essential to distinguish between “difference” and “disparity” because remedies and policy implications for each of these are different. Nonetheless, documenting both differences and disparities is important.

The authors report that among those eligible, 11.5% of white patients and 4.7% of black patients were referred for cardiac rehabilitation. These data suggest that rehabilitation services are grossly underused. Reasons for underuse may include lack of knowledge about the benefits of rehabilitation among primary care and other referring physicians. The rehabilitation academy should consider active steps to educate referring physicians about the benefits of rehabilitation. In addition, it is critical to better understand the reasons for differential use of cardiac rehabilitation by race.

Nitin B. Jain, MD, MSPH
Department of Medicine
Newton-Wellesley Hospital
Newton, Massachusetts
Channing Laboratory
Brigham and Women’s Hospital and Harvard Medical School
Boston, Massachusetts

Jeffrey N. Katz, MD, MS
Section of Clinical Sciences and Center for Outcomes Research
Division of Rheumatology, Immunology, and Allergy, and
Department of Orthopaedic Surgery
Brigham and Women’s Hospital
Boston, Massachusetts

REFERENCES

RE: RACIAL DISPARITIES IN ACCESS TO CARDIAC REHABILITATION

The Authors Respond: We thank Drs. Jain and Katz for their careful consideration of our article, “Racial Disparities in Access to Cardiac Rehabilitation.”1 Jain and Katz use the five formal criteria of Rathore and Krumholz2 to justify the use of the term “disparity,” arguing that our study has failed to meet at least one of the requirements. Although it has not been established that these criteria have been widely adopted, we maintain that Drs. Jain and Katz are inaccurate in their critique. In fact, we argue that...
on the basis of these five criteria, our study does define a case of differential use of services by race as a disparity.

Jain and Katz agree that we satisfy criteria 1, 2, 4, and 5. We satisfy the first criterion (Patients eligible for the test or procedure are being evaluated) by using the American College of Cardiology/American Heart Association guidelines for candidates for cardiac rehabilitation to define our study population. We satisfy the second criterion (The analysis accounts for treatment contraindications) because the study population was taken from similar circulatory diagnoses.

The fourth criterion (There has been risk adjustment for patient factors including demographic, clinical, and social characteristics) is met because the multivariate analyses were adjusted for social and demographic factors as well as clinical factors such as receipt of cardiac procedures. Jain and Katz argue that the attenuation of the odds ratio from 2.5 to 1.8 indicates that the effect of race may be mediated by other variables such as insurance status and receipt of specialty consultation. We would, however, caution against this overinterpretation of multivariate adjustment. A person’s race influences insurance before exposure to measured social, physiologic, and psychological status. All of these factors are inherently related to race in a racially stratified society. So, it is not possible to identify root cause, because all the variables are interrelated.

We meet the fifth criterion (Racial variations in treatment are associated with poorer patient outcomes including progression of disease) because prior studies have shown that participation in cardiac rehabilitation reduces the risk of ventricular arrhythmias and sudden death and reduces mortality by 20% in comparison with nonparticipants.5–6

Jain and Katz claim that we do not meet criterion 3 (Have patient preferences been considered?) by calling attention to the role of patient preferences. We cite that this study evaluated referral for cardiac rehabilitation. As we show in Table 1, among patients who were referred for cardiac rehabilitation, 80% of blacks received cardiac rehabilitation and 69% of whites received cardiac rehabilitation. There is, therefore, no evidence of disparities in patient preference for use of cardiac rehabilitation. And, in fact, the African American patients in this study seem to be more likely than whites to adhere to doctor recommendations regarding cardiac rehabilitation. Additionally, in analyses of these data reported in another article, we found that among African American patients who were referred for other procedures, there was no race difference in acceptance.7

Critics of healthcare-disparities research often site patient preferences as a source of disparities; however, there is little research to support this claim. The problem of health disparities, it seems, has more to do with race differences in obtaining a referral. As the Institute of Medicine concluded in the 2002 report Unequal Treatment, A small number of studies suggest that racial and ethnic minority patients are more likely than white patients to refuse treatment. These studies find that differences in refusal rates are generally small and that minority patient refusal does not fully explain healthcare disparities.7 (p 8)

Although we did not specifically assess preferences for cardiac rehabilitation, the available evidence suggests that patient preferences are unlikely to be a significant source of race differences in use of cardiac rehabilitation in our study. Because all five criteria were met, we have correctly used the term “disparity.” This study does not show differences in use of services, but it shows differences in referral rates between blacks and whites, with similar use among those referred. On the basis of the Rathore and Krumholz8 criteria, this study reflects the differential provision of appropriate care to black patients after controlling for health system factors. We would also argue that this study meets the criteria for disparity according to the Institute of Medicine definition in Unequal Treatment, which defines disparity as racial or ethnic differences in the quality of health care that are not attributable to access-related factors or clinical needs, preferences, or appropriateness of intervention.8

Most importantly, this study highlights the inequitable delivery of health care that leads to poorer outcomes among the disadvantaged black population.9 Different race groups have different levels of exposure to health risks that result in poorer health outcomes.

Finally, we would echo the fact that cardiac rehabilitation is grossly underused among those who are eligible, without regard to race. We would further add that there is also an underlying racial disparity in the referral of patients for cardiac rehabilitation. Efforts should be instituted to educate referring physicians about the benefits of cardiac rehabilitation and to identify interventions to eliminate the racial disparity.

Patricia C. Gregory, MD
Department of Physical Medicine and Rehabilitation
The University of North Carolina–Chapel Hill
Chapel Hill, North Carolina

Thomas A. LaVeist, PhD
Center for Health Disparities Solutions
The Johns Hopkins Medical Institutions
Baltimore, Maryland

Crystal F. Simpson, MD, MHS
Division of Geriatric Medicine
The Johns Hopkins Medical Institutions
Baltimore, Maryland

REFERENCES
4. Kaufman JS, Cooper RS: Commentary: considerations for use of racial/


A Technique for Ultrasound-Guided Intrathecal Drug-Delivery System Refills

Mark-Friedrich B. Hurdle, MD, Adam J. Locketz, MD, and Jay Smith, MD

From the Department of Anesthesia Pain Medicine (M-FBH, AIL) and Department of Sports Medicine (JS), Physical Medicine and Rehabilitation, Mayo Clinic College of Medicine, Rochester, Minnesota.

A 41-yr-old woman with kyphoscoliosis received an intrathecal drug delivery system (ITDDS) for control of pain and myoclonus. Unfortunately, over the course of several years, her posture worsened and her body mass increased significantly. Her ITDDS pump reservoir (Medtronic Neurological, Minneapolis, MN) became impossible to consistently access without fluoroscopic guidance. As the result of her kyphoscoliosis and pain with spinal extension, positioning on the fluoroscopic table required the maximal assistance of three people for both the transfer and positioning. Consequently, we elected to image and fill the ITDDS using ultrasound (US) guidance via the technique described herein.

The skin is palpated to localize the subcutaneous position of the pump. Using an US machine (Diagnostic Instruments, Inc., Annapolis, MD) and gel, the pump is visualized with a 12-MHz linear-array probe. The hyperechoic signal of the metallic pump, with posterior acoustic shadowing, can be clearly seen. The optimal imaging angle can then be obtained by rocking the transducer along both the transverse and longitudinal axes, until the intensity of the signal is maximized. Next, one can sweep the transducer across the pump until the hypoechoic, centrally located port is visualized (Figure 1). The port is centered in the scan image and then a 22-gauge noncoring needle is inserted into the skin perpendicular to the long axis of the US transducer (i.e., a short-axis approach). Once the needle tip enters the imaging field, the operator can visualize the tip. The operator should repeatedly angulate and advance the needle until it abuts the port, being careful not to allow the needle tip to extend beyond the imaging field. The operator can also appreciate port contact with the needle tip via tactile feel. Once the port is contacted, the needle is advanced, and the medication delivered. Simultaneous real-time US during the injection can confirm that no medication collects superficial to the ITDDS, thus confirming proper deposition of medication.

Increasingly, patients with refractory pain and spasticity are receiving ITDDS. Care providers who refill pump reservoirs with any intrathecal medication must be knowledgeable about the life-threatening consequences associated with refilling errors. The literature contains numerous case reports of overdose from ITDDS for a variety of reasons. Although not formally studied to date, the incidence of complications will likely increase as the prevalence of ITDDS increases.

The use of US technology for imaging is becoming more widespread in the clinical setting. Because of the technical training required to master medical US technology, reluctance may exist on the part of operators to use US for guiding an ITDDS refill. However, US-guided ITDDS is both feasible and practical. We have found a successful and simple technique for properly positioning the needle tip over the refill port to facilitate ITDDS access for any reason. To our knowledge, this is the first description of this or any such technique in the medical literature.

We have successfully refilled multiple hard-to-access pump reservoirs using this technique. No complications have occurred, and both patients and operators have reported a universally high acceptance of the procedure. US imaging allows both the patient and operator to avoid the ionizing radiation of fluoroscopy and reduces the need for multiple needle passes in difficult cases, saving time and reducing patient discomfort. In addition, the soft tissue imaging pro-
vided by US can detect potentially clinically significant fluid collections such as a seroma, hematoma, or abscess. Using US guidance seems to be a safe and effective way to assist with refilling hard-to-access ITDDS pump reservoirs.

REFERENCES


