Editorial Notes to Begin 2005.
Frontera, Walter R. MD, PhD

Clinical Trial Registration: A Statement from the International Committee of Medical Journal Editors.
De Angelis, Catherine MD, MPH; Drazen, Jeffrey M. MD; Frizelle, Frank A. MB, ChB, MMedSc., FRACS; Haug, Charlotte MD, PhD, MSc; Hoey, John MD; Horton, Richard FRCP; Kotzin, Sheldon MLS; Laine, Christine MD, MPH; Marusic, Ana MD, PhD; Overbeke, A John P. M. MD, PhD; Schroeder, Torben V. MD, DMSc; Sox, Hal C. MD; Van Der Weyden, Martin B. MD

Physicians with Disabilities and the Physician Workforce: A Need to Reassess Our Policies.
DeLisa, Joel A. MD, M.S; Thomas, Peter Esq

Influence of the Prospective Payment System on Speech-Language Pathology Services.
Frymark, Tobi B. MA; Mullen, Robert C. MPH

Burns, Stephen P. MD; Breuninger, Amy MPT; Kaplan, Carri MPT; Marin, Heather MPT

Unintentional Vascular Uptake in Fluoroscopically Guided, Contrast-Confirmed Spinal Injections: A 1-Yr Clinical Experience and Discussion of Findings.
Stretanski, Michael F. DO; Chopko, Bohdan PhD, MD

Habitual Physical Activity Levels Are Associated with Biomechanical Walking Economy in Children with Cerebral Palsy.
Maltais, Desiree B. MSc; Piernyowsk, Michael R. PhD; Galea, Victoria A. PhD; Matsuzaka, Akira MSc; Bar-Or, Oded MD

Predicting Outcomes after Hip Fracture Repair.
Kagaya, Hitoshi MD, DMSc; Takahashi, Hitomi RPT; Sugawara, Keiyu RPT; Dobashi, Mayumi RPT; Kiyokawa, Noritaka RPT; Ebina, Hazuki RPT

Influence of Aquatic and Weight-Bearing Exercises on Quantitative Ultrasound Variables in Postmenopausal Women.
Ay, Alev MD; Yurtkuran, Merih MD

Functional Neuroimaging and Cognitive Rehabilitation for People with Traumatic Brain Injury.
Strangman, Gary PhD; O'Neil-Pirozzi, Therese M. PhD; Burke, David MD; Cristina, Dana BS; Goldstein, Richard PhD; Rauch, Scott L. MD; Savage, Cary R. PhD; Glenn, Mel B. MD
Case Report
Injections

Transforaminal Cervical Blood Patch for the Treatment of Post-Dural Puncture Headache.
Slipman, Curtis W. MD; El Abd, Omar H. MD; Bhargava, Amit MD; DePalma, Michael J. MD; Chin, Kingsley R. MD

Visual Vignette

Ultrasound-Guided Aspiration of Symptomatic Rotator Cuff Calcific Tendonitis.
Cooper, Grant MD; Lutz, Gregory E. MD; Adler, Ronald S. MD
EDITORIAL

Editorial Notes to Begin 2005

With this first issue of the year 2005, we celebrate the 83rd anniversary of the American Journal of Physical Medicine and Rehabilitation (AJPMR). AJPMR was founded in 1922 and has been published since the very first issue by Williams & Wilkins (now Lippincott Williams & Wilkins). AJPMR is a vehicle for communicating new ideas and research findings about the scientific foundation and the practice of physical medicine and rehabilitation. We continue to work to make this possible. During the calendar year 2004, the number of manuscripts submitted to the Journal for review continued to increase (Fig. 1), evidence that the scientific activity in the field, both nationally and internationally, is growing significantly. We want the reader and his or her patients and students to benefit from this activity.

One of the objectives of the Editorial Board is to make sure that authors, reviewers, and readers find their experiences with the Journal a learning opportunity and a professionally rewarding investment of time. To facilitate the achievement of this goal, we introduce with this issue of the Journal a new internal page design with two columns and a new format for tables and figures. During the last few months, we have used advanced information technology to make both the submission of articles and the review process easier and more efficient. Further, in an attempt to reward the reviewers for their time and expertise, we have added a new feature that makes CME credits available online to them. Finally, the on-line version of the Journal continues to thrive, with more users depending on the electronic access for their articles; a general trend in medical and scientific publishing.

With this issue of the Journal, we are adding a special “Invited Review” section. Experts in very specific areas (“pharmacological management of agitation in patients with traumatic brain injury” as opposed to “brain injury rehabilitation”) related to our field will be invited to write an in-depth critical review of selected topics of interest to rehabilitation professionals. The reviews will address the scientific basis and the clinical practice aspects of a particular topic. The range of topics include functional magnetic resonance imaging and brain injury, knee injuries in women athletes, pregnancy and musculoskeletal pain, hypertrophic scarring in burns, and many others. We plan to publish this section, like the very successful “Visual Vignettes” section, on a regular monthly schedule. Our Editor for review articles, Dr. Henry Lew, has been working diligently to put together a series that I am sure will be of interest to many readers.

Some changes in the composition of the Editorial Board must be noted. I would like to thank Dr. John Redford, who is finishing his tenure as a member of the Board, for his invaluable service to the Journal. Dr. Redford provided excellent advice on the quality and acceptability of many scientific manuscripts, especially in the area of orthotics. Also, Dr. John Baker, who has provided excellent reviews and advice in the area of outcomes research, ends his participation in the Board.

A current member of the Editorial Board, Dr. Carl Granger, a nationally and internationally recognized expert in the field of functional assessment, has been appointed Associate Editor. Three distinguished members of our rehabilitation medicine community have been appointed as new members of the Editorial Board: Dr. Curtis Slipman, Associate Professor of Physical Medicine and Rehabilitation and Director of the Spine Center at the University of Pennsylvania, Dr. Peter Esselman, Associate Professor of Physical Medicine and Rehabilitation in the Department of Rehabilitation Medicine at the University of Washington and Chief of Rehabilitation Medicine at the Harborview Medical Center in Seattle, and Dr. Martin K. C. Childers, Associate Professor of Physical Medicine and Rehabilitation at the University of Missouri. Finally, Dr. Neville Hogan, Professor of Mechanical Engineering and Brain and Cognitive Sciences at the Massachusetts Institute of Technology, joins the Board as an Associate Editor. I wish to extend a most sincere welcome to all of them.

Research Summit

A research summit titled “Building Research Capacity in Rehabilitation Medicine” will take place in Washington, DC, on April 28–29, 2005. The summit is organized and sponsored by the Foundation for Physical Medicine and Rehabilitation, the Association of Academic Physiatrists, the American Academy of Physical Medicine and Rehabilitation, and the American Congress of Rehabilitation Medicine. It should not escape our attention that these organizations have enthusiastically agreed to work together to address an important issue that could shape the future of rehabilitation medicine. Further, rehabilitation professions such as occupational therapy, rehabilitation engineering, physical therapy, and many others have also been invited to participate. The goals of the Summit are to stimulate discussion about research capacity, to disseminate in the form of several publications the most impor
tant aspects of the discussion, and to develop an action plan that could be adopted and supported by all professional organizations in the field and by funding agencies, foundations, and institutions.

The Program Committee has chosen five elements of research capacity to provide a framework for the discussion. The program will present a critical analysis of the current status, the identification of weaknesses, and the definition of an active plan to enhance research capacity in the field. The five elements that will guide the discussion are: (1) researchers (training, mentoring, retention), (2) infrastructure and research environment (institutional culture and resources), (3) funding (sources and support for research capacity), (4) partnerships and collaborations (other fields of scientific inquiry), and (5) metrics (how to measure research capacity).

The program will include, among other things, keynote lectures by Harvey Fineberg, MD (President of the Institute of Medicine) and Jordan Cohen, MD (President of the AAMC), paper presentations on the five elements of research capacity mentioned above, reactions to the papers by distinguished researchers in the field, small group discussions, and a plenary session for deliberations and identification of proposals for an action plan. In future issues, we hope to bring to you in the pages of this Journal the results of this very important meeting.

**Registry of Clinical Trials**

Recently, a group of editors of 11 of the leading medical journals in the world announced their decision to publish only the results of clinical trials that were registered in a public registry when the trial began. This decision will apply to any trial starting subject enrollment after July 1, 2005. One of several reasons for this very significant action is the problem of selective reporting in drug trials that may result in the publication of positive results and the suppression of negative results that are then kept from the scrutiny of the scientific and medical communities. In this issue of the *Journal*, we reproduce this article because this registry could change noticeably the conduct of a very important type of research that rehabilitation professionals are being asked to do more often. An article in *Science* reports that this could result in legislation that could mandate the creation of such a registry to: (1) register all United States drug trials at their launch, (2) list eligibility requirements for participants, (3) list funding sources, (4) post results including those not published in journals, and (5) fine noncompliant trial sponsors. Our *Journal* will follow closely this development and keep our readers and potential authors informed.

Walter R. Frontera, MD, PhD
Editor-in-Chief

**REFERENCES**


2. Couzin J: Legislators propose a registry to track clinical trials from start to finish. *Science* 2004;305:1695
Altruism and trust lie at the heart of research on human subjects. Altruistic individuals volunteer for research because they trust that their participation will contribute to improved health for others and that researchers will minimize risks to participants. In return for the altruism and trust that make clinical research possible, the research enterprise has an obligation to conduct research ethically and to report it honestly. Honest reporting begins with revealing the existence of all clinical studies, even those that reflect unfavorably on a research sponsor’s product.

Unfortunately, selective reporting of trials does occur, and it distorts the body of evidence available for clinical decision-making. Researchers (and journal editors) are generally most enthusiastic about the publication of trials that show either a large effect of a new treatment (positive trials) or equivalence of two approaches to treatment (noninferiority trials). Researchers (and journals) typically are less excited about trials that show that a new treatment is inferior to standard treatment (negative trials) and even less interested in trials that are neither clearly positive nor clearly negative, since inconclusive trials will not in themselves change practice. Irrespective of their scientific interest, trial results that place financial interests at risk are particularly likely to remain unpublished and hidden from public view. The interests of the sponsor or authors notwithstanding, anyone should be able to learn of any trial’s existence and its important characteristics.

The case against selective reporting is particularly compelling for research that tests interventions that could enter mainstream clinical practice. Rather than a single trial, it is usually a body of evidence, consisting of many studies, that changes medical practice. When research sponsors or investigators conceal the presence of selected trials, these studies cannot influence the thinking of patients, clinicians, other researchers, and experts who write practice guidelines or decide on insurance-coverage policy. If all trials are registered in a public repository at their inception, every trial’s existence is part of the public record and the many stakeholders in clinical research can explore the full range of clinical evidence. We are far from this ideal at present, since trial registration is largely voluntary, registry data sets and public access to them varies, and registries contain only a small proportion of trials. In this editorial, published simultaneously in all member journals, the International Committee of Medical Journal Editors (ICMJE) proposes comprehensive trials registration as a solution to the problem of selective awareness and announces that all eleven ICMJE member journals will adopt a trials-registration policy to promote this goal.

The ICMJE member journals will require, as a condition of consideration for publication, registration in a public trials registry. Trials must register at or before the onset of patient enrollment. This policy applies to any clinical trial starting enrollment after July 1, 2005. For trials that began enrollment before this date, the ICMJE member journals will require registration by September 13, 2005, before considering the trial for publication. We speak only for ourselves, but we encourage editors of other biomedical journals to adopt similar policies. For this purpose, the ICMJE defines a clinical trial as any research project that prospectively assigns human subjects to intervention or comparison groups to study the cause-and-effect relationship between a medical intervention and a health outcome. Studies designed for other purposes, such as to study pharmacokinetics or major toxicity (e.g., phase 1 trials), would be exempt.

The ICMJE does not advocate one particular registry, but its member journals will require authors to register their trial in a registry that meets several criteria. The registry must be accessible to the public at no charge. It must be open to all prospective registrants and managed by a not-for-profit organization. There must be a mechanism to ensure the validity of the registration data, and the registry should be electronically searchable. An acceptable registry must include at minimum the following information: a unique identifying number, a statement of the intervention (or interventions) and comparison (or comparisons) studied, a statement of the study hypothesis, definitions of the primary and secondary outcome measures, eligibility criteria, key trial dates (registration date, anticipated or actual start date, anticipated or actual date of last follow-up, planned or actual date of closure to data entry, and date trial data considered complete), target number of subjects, funding source, and contact information for the principal investigator. To our knowledge, at present, only www.clinicaltrials.gov, sponsored by the United States National Library of Medicine, meets these criteria.
requirements; there may be other registries, now or in the future, that meet all these requirements.

Registration is only part of the means to an end; that end is full transparency with respect to performance and reporting of clinical trials. Research sponsors may argue that public registration of clinical trials will result in unnecessary bureaucratic delays and destroy their competitive edge by allowing competitors full access to their research plans. We argue that enhanced public confidence in the research enterprise will compensate for the costs of full disclosure. Patients who volunteer to participate in clinical trials deserve to know that their contribution to improving human health will be available to inform health care decisions. The knowledge made possible by their collective altruism must be accessible to everyone. Required trial registration will advance this goal.

Catherine De Angelis, MD, MPH
Editor-in-Chief, JAMA

Jeffrey M. Drazen, MD
Editor-in-Chief, New England Journal of Medicine

Frank A. Frizelle, MB, ChB, MMedSc., FRACS
Editor, The New Zealand Medical Journal

Charlotte Haug, MD, PhD, MSc
Editor-in-Chief, Norwegian Medical Journal

John Hoey, MD
Editor, CMAJ

Richard Horton, FRCP
Editor, The Lancet

Sheldon Kotzin, MLS
Executive Editor, MEDLINE National Library of Medicine

Christine Laine, MD, MPH
Senior Deputy Editor, Annals of Internal Medicine

Ana Marusic, MD, PhD
Editor, Croatian Medical Journal

A. John P. M. Overbeke, MD, PhD
Executive Editor, Nederlands Tijdschrift voor Geneeskunde (Dutch Journal of Medicine)

Torben V. Schroeder, MD, DMSc
Editor, Journal of the Danish Medical Association

Hal C. Sox, MD
Editor, Annals of Internal Medicine

Martin B. Van Der Weyden, MD
Editor, The Medical Journal of Australia

This statement has been previously published. Copyright 2004 Massachusetts Medical Society. Reprinted with permission.
Physicians with Disabilities and the Physician Workforce: A Need to Reassess Our Policies


People with disabilities make up about 20% of the population, yet only a tiny fraction of matriculants to medical school have disabilities. Attempts to define core technical standards and competencies have not kept pace with technological changes, diverse specialization, and changing practice options. This has resulted in the inappropriate exclusion of some people with disabilities. Medical schools determine how any qualified applicant, regardless of physical or cognitive ability, can be effectively accommodated and counseled in achieving the most appropriate medical career. A serious effort to redefine the technical standards and core competencies of the 21st century medical education at the undergraduate and graduate levels would likely resolve many of the troubling questions regarding medical students with disabilities. We have made some recommendations to organized medicine for constructing an agenda to address these issues.

Key Words: Disabilities, Medical Education, Medical Student, Accommodations, Competencies

INTRODUCTION

In 1997, Dr. Jordan Cohen, the president of the Association of American Medical Colleges, issued a moral charge to the medical profession: “to take active steps to ensure that our healthcare practitioner community mirrors society’s gender, racial, and ethnic mix.” Why? Because it is a matter of social justice and equality; because it is a means to improve access to health care on the part of the underserved; because it is a way to deliver culturally competent care—particularly to minority populations who are often disproportionately affected by healthcare problems; and because it just makes sense to fully use the rich and diverse pool of our nation’s people to better manage the healthcare system. Recently, Dr. Cohen has expanded the scope of this issue beyond considerations of race, ethnicity, and gender to include issues of disability. His editorial entitled “Reconsidering ‘Disabled’ Applicants” challenges medicine to reconsider what it takes to be a capable doctor.

We think Dr. Cohen has it right, and we firmly support and applaud his efforts to “bridge the diversity gap” in our healthcare training programs. In particular, we
would like to elaborate on the underrepresentation of people with disabilities who are trained as physicians. Although people with disabilities make up about 20% of the population, only a tiny fraction of matriculants to medical school have disabilities. Though data are limited, best estimates indicate that people with physical disabilities comprise less than 1% of medical school graduates. The percentage of physicians with disabilities in practice is higher, with estimates ranging from 2 to 10%, suggesting that although getting into medical school is a hurdle for people with disabilities, to the entry point it seems there is a stronger commitment to keep physicians in training or in practice. If disability occurs subsequently, and the age-specific prevalence of major chronic conditions remains unchanged, the absolute number of Americans with functional limitations is expected to rise by more than 300% by 2049. If we interpolate this data to practicing healthcare providers, we can expect that the number of physicians in practice who become disabled also will increase.

So we are left with the question: why are so few people with disabilities physicians? Is it a lack of ability? Is it a lack of opportunity? Perhaps this argument may have been compelling in the days before diverse specialization, changing practice options, and technological advances, but it is less so today. Attempts to define core technical standards and competencies have not always kept up with these changes, and have resulted in the inappropriate exclusion of some people with disabilities. Given increasing technological resources and the changing scope of medical practice, it is imperative that we embrace these issues head-on. Yet serious attempts by our healthcare training programs to move these conceptual issues forward have not progressed significantly since the July 26, 1990 passage of the Americans with Disabilities Act (ADA), or the 1993 passage of the Rehabilitation Act, both of which established disability as the seventh protected class under federal nondiscrimination law. Individual medical schools have been left to interpret and apply the law, and some schools have done a better job than others in attempting to address this issue. Little guidance has been issued to address critical questions such as the core technical standards and skills that all medical students must possess to meet the demands of current and future medical practice, and what constitutes reasonable modification or reasonable accommodations for individuals with disabilities.

The remainder of this paper will present arguments in favor of a serious effort to include people with disabilities in medical school and postgraduate training programs, as well as more organized strategies for keeping physicians in the profession after disability occurs. In fact, a serious effort to redefine the technical standards and core competencies of 21st century medical education at both the undergraduate and graduate levels would likely resolve many of the troubling questions regarding medical students with disabilities. We will conclude with some concrete recommendations for constructing an agenda to address these goals.

**Definition of Disability**

Disability conditions are diverse in their causes, nature, timing, pace, and societal implications. Some are congenital, others are acquired. Some occur suddenly with injury or accident; others arise slowly, with progressive debility. There are various categories of disability, including physical, mental, sensory (vision and hearing deficits), and developmental disability, each of which impacts a significant portion of the population. It is likely that if you live long enough, you will experience disability at some point in your lifetime. Disability is a widespread phenomenon, and represents a minority group that everyone is at risk to join. Although there is no single consensus definition of disability, (Table 1), for the purposes of this paper we will reference the definition of disability in the ADA(i.e., a “physical or mental impairment that substantially limits one or more of the major life activities”).

**What Does It Mean To Be a Physician in 2004?**

The medical profession has arguably seen more changes in the last 50 yr than during the preceding millennium. No longer dominated by primary care solo practitioners, we have become a nation of specialists (about 70% of the practicing physicians work force), highly reliant on technological resources and often working in teams or networks. Nuanced physical examination techniques are being displaced by MRIs and echocardiograms that offer greater precision. Nurse practitioners and physician assistants are playing increasingly important professional and supportive roles within healthcare teams. The medical database is exploding with information defying physicians to keep up, let alone assimilate a morass of complicated and often contradictory studies in an effort to make evidence-based decisions. One of the biggest changes in health care these days is that patients spend much less time in hospitals compared with years past. As such, medical schools have faced considerable difficulty reorienting traditionally inpatient-based training to the new outpatient reality. Much of this world is being dominated by technology and automation.

Medical schools and resident training programs are scrambling to keep pace with these changes.
Though the core mission of medical schools and training programs has not changed—to train effective, competent and compassionate physicians best able to serve the needs of society—the strategies for achieving these goals have changed. One has only to review curricular changes over the last two decades to appreciate that critical thinking and communication skills are receiving greater emphasis, whereas technical skills and rote memorization have declined in curricular emphasis.15,16

These tensions are far from resolved. One can detect real ambivalence, or, perhaps more accurately, uncertainty about the essential requirements for graduating a physician from medical school. In the absence of a crystallized consensus on this topic, the written standards and guidelines seem somewhat inert and dated, with change occurring slowly. In 1979, an Association of American Medical Colleges (AAMC) advisory panel recommended technical standards to guide admission to medical school. These have been defined as the essential functions for a graduating medical student.

The panel concluded that a candidate for the MD degree must have abilities and skills in the following areas:17

1. Observation—performed in a reasonably independent manner
2. Communication skills
3. Motor skills—performed in a reasonably independent manner
4. Intellectual-conceptual, integrative, and qualitative abilities
5. Behavioral and social attributes

After the passage of the ADA, the AAMC published a follow-up document on medical school admission requirements in the United States and Canada (1991–1992)18 It states that “candidates for the MD degree must have somatic sensation and the functional use of senses of vision and hearing.” It further states that a candidate’s diagnostic skills will also be lessened without the functional use of the senses of equilibrium, smell and taste. Additionally, students must have sufficient exteroceptive sense (touch, pain, and temperature), sufficient proprioceptive sense (position, pressure, movement, stereognosis, and vibratory), and sufficient motor function to carry out activities “necessary for education of the physician.” They must also be able to consistently, quickly, and accurately integrate all information received by whatever sense(s) employed, and they must have the intellectual ability to learn, integrate, analyze, and synthesize these data.18 It would seem that this document did not address psychiatric disability, or learning disabilities.

Though this report was not AAMC policy, it was intended as a guideline for medical schools to use in establishing their own technical standards, indicating that the faculty of each medical school must review its own curriculum and reflect on its own educational goals.18 The report indicated that schools were to provide reasonable accommodations, but it was vague in defining “reasonable.” Ironically, over the years these principles seem to have had the effect of preventing talented individuals with disabilities from attending American medical schools.17


<table>
<thead>
<tr>
<th>TABLE 1 Definitions of Disability</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Social Security Administration</strong></td>
</tr>
<tr>
<td>“The inability to engage in any substantial gainful activity by reasons of any medically determinable physical or mental impairments which can be expected to result in death or which has lasted or can be expected to last for a continuous period of not less than 12 months (defined in terms of functional limitations as they effect employability).”8</td>
</tr>
<tr>
<td><strong>Americans with Disabilities Act (ADA): Section 3</strong></td>
</tr>
<tr>
<td>“Someone who has: (A) physical or mental impairment that substantially limits one or more of the major life activities; . . . (B) a record of such impairment; or (C) being regarded as having such an impairment.” ADA defines disability from the perspective of physical or mental impairments, but also recognizes barriers, or the failure to provide reasonable accommodations, can give to the denial, or limitation of opportunities.”9</td>
</tr>
<tr>
<td><strong>World Health Organization</strong></td>
</tr>
<tr>
<td>Disability is an “umbrella term for impairments, activity limitations, or participant restrictions” (p. 3); “a person’s functioning and disability (represent) a dynamic interaction between health conditions (diseases, disorders, injuries, traumas, etc.) and contextual factors,” including environmental, social, and personal attributes. (p. 8). This places disability within a broad “biopsychosocial” perspective, integrating the medical and social models.10</td>
</tr>
<tr>
<td><strong>Medical Definition:</strong></td>
</tr>
<tr>
<td>The physical disadvantage that results from impairment; the difficulty in performing physical tasks. The individual requires rehabilitation and possibly accommodation to function as well as others.17</td>
</tr>
</tbody>
</table>
DeLisa and Thomas

Am. J. Phys. Med. Rehabil. • Vol. 84, No. 1

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

The Thorny Issue of Technical Standards

The most recent effort to define appropriate technical standards is contained within the initial 1999 report of the Medical Schools Objectives Project. This report organized medical education goals and objectives into the categories of altruism, knowledge, skills, and sense of duty. Under the heading “Physicians Must Be Skilful,” the report outlines technical standards in much the same fashion as the 1979 AAMC report, requiring that graduates be able to perform a complete physical examination, perform and interpret diagnostic tests, and respond appropriately to immediately life-threatening medical conditions. The Medical School Objectives Project report retains an emphasis on physical technical performance, demanding that graduates demonstrate proficiency in such skills as venipuncture, lumbar puncture, and suturing lacerations.19

Although discussions of medical students with disabilities focus primarily on technical standards, many question the validity of a strong emphasis on technical skills. Reichgott asked several important questions regarding the role of technical skills and their relative importance when compared with other requirements of graduates, such as knowledge/intelligence, professional attitude, and the ability to communicate and interact effectively. For example, “Is the hands-on, personal touching experience afforded by the course in physical diagnosis necessary for the effective integration of basic science knowledge and the understanding of pathophysiology? If a trained assistant does the physical exam and provides data to the student... does this really impose a negative ‘interpreter’ effect?”20 Given the diversity of available specialties, one must even question whether there should be any mandatory physical technical skills in medical school.

Van Matre et al. sent a 3-page questionnaire to faculty, residents and third-year medical students affiliated with Northwestern University’s Feinberg School of Medicine. The majority of the survey respondents, regardless of level of training or disability status, believed that disabilities affecting motor skills are less likely to impede the practice of medicine than those that affect the ability to observe or communicate. Technical skills used in interpretation and observation, such as palpation and percussion, were more important to respondents than those that are more procedural, such as inserting an intravenous catheter or tying sutures.21

Undifferentiated Graduate vs. Undifferentiated Curriculum

Medical schools have almost exclusively enrolled those students who seem to have the potential to enter any existing field of medicine; these are considered to be the undifferentiated graduate. However, significant differentiation of physicians into various specialties and subspecialties can serve as an argument for less rigidity in demanding that all students demonstrate competence in procedures that are not relevant to their future expected practices. Medical specialization has segmented the physician’s workforce from a more homogenous group to one concentrating on specific body systems or disease entities.

Healthcare professionals adequately trained for the future will need to know what informational resources to use; how to gather necessary data; how to integrate complex information, make diagnoses, and develop treatment plans; and how to effectively use changing technological resources, work with teams, and communicate with diverse populations. These skills are largely cognitive and not physical, raising questions about the adequacy of the current approach to medical training. Although applicants with disabilities should meet the same cognitive admission standards as their peers without disabilities, these standards need to be fair and reflective of the essential criteria for the profession, and not serve as a barrier or deterrent to otherwise qualified applicants with disabilities.7 Medical schools need to answer several questions, including what it means to be a doctor today, what constitutes good doctoring, and what are the truly nonnegotiable elements comprising a basic medical education.2

However, we absolutely oppose a tracking system, where an individual is admitted to medical school under the presumption that he or she will be designated to a specific postgraduate specialty. Each student must be handled on a case-by-case basis.

Principles in the Training of a Physician with a Disability

Although most medical students take similar courses and clerkships through the first three years, by the fourth year students focus increasingly on their own particular interests. Medical students do not graduate as pluripotent physicians. Each field (specialty) requires different psychological, emotional, verbal, intellectual, and technical skills. Medical school provides students with exposure to help target their eventual practice. Also, over time most practitioners gain new knowledge and skills in focused activities of choice, and are less concerned about knowing all of medicine, or about being skilled in all diagnostic and therapeutic procedures.20
There Are Two Overlying Principles That Must Be Adhered to While Training a Physician With a Disability: Protection of Patients

The person with disabilities (trainee) has the ability to practice at a level comparable with that of the person without impairment. Patient well-being is held sacrosanct by the Hippocratic tradition to “do no harm.” The achievement of technical standards appropriate to the type of practice the trainee proposes to pursue is essential (M.G. Stineman, personal communication).

Rights of the Trainee

It is essential to respect the creative solutions that people with disabilities often employ to perform tasks in alternate ways. The ability to perform the task at a defined level of quality should be emphasized rather than the process by which the task is accomplished. We need to be flexible and consider what is possible through hard work and low or high technology (M.G. Stineman, personal communication).

Arguments Against Training People with Disabilities in Medical Education

The health professions have strong societal fiduciary responsibilities that include the protection of patients and the wise use of resources. As part of a “social contract” of sorts, with tax dollars used to supplement the education of physicians, medical schools have a responsibility to ensure that they are training physicians who will be able to best meet the needs of society. Such arguments have been used in the past to counter why more women should not be admitted to medical schools (i.e., there is less value in return for the dollar, given that women historically have not worked as many hours or years as their male counterparts). Given that medical school gender enrollment is now about equal, this gender bias is obviously outdated, and no longer operative. Indeed, the benefits of having women in all aspects of the health professions have been realized. Similar arguments are currently being offered regarding physicians with disabilities.

Arguments in Favor of People with Disabilities in Medical Education

Physicians with disabilities may bring to their practice unique perspectives and empathy because of their personal experience with disability. Misinformation and prejudice about disability abound in the health professions and, indeed, are often perpetuated by healthcare providers. In numerous studies, healthcare professionals have been known to be more negative in their estimates of the quality-of-life of a person with a disability than the person with the disability him or herself. These negative attitudes can have an effect on the framing of information and the very treatments offered. Yet little time is spent in medical school curricula around issues of disability, despite the fact that practicing clinicians will invariably come into contact with and/or treat a substantial number of people with disabilities during their career. Mutual respect between doctors and patients with disabilities, regardless of how severe the disability may be, seems to be one of the most effective ways to break down barriers and dispel prejudices.

A Survey of the Case Law

A survey of the case law that has developed since passage of the ADA and Rehabilitation Act reveals that these nondiscrimination laws have been used in the medical school context in three primary ways: requests for accommodations in taking examinations, primarily due to various forms of learning disabilities; challenges of denials of admission to medical school based on disability; and challenges to dismissal from medical school based on disability. Although a number of cases have been decided in favor of applicants and students with disabilities, the vast majority of cases do not grant relief to people with disabilities, primarily because it is difficult to prove that the person’s disability was the cause of the dismissal, admission rejection, or failure of an examination. This places the burden of meaningfully addressing the lack of medical school applicants and students with disabilities on the medical schools themselves.

Recommendations

The need for program modifications and reasonable accommodations differs for students, residents, and faculty. A student’s focus is on educational requirements and on meeting the diverse demands of the basic sciences and clinical years. Faculty members with disabilities can tailor their practices to minimize the need for accommodations. Residency, however, is truly a mixture of service and education, and offers perhaps the greatest challenges in terms of disability considerations. Residents may need to perform in certain services essential to the residency program that would not be necessary for students or faculty members. Even within a specialty, not all programs have the same “service” requirements. Should this work obligation be a barrier to satisfactory completion of residency training? Meier as well as Hartman and Hartman have questioned why an applicant should be denied a chance to practice medicine just because he or she cannot perform certain procedures required of the specialties.
Medical schools determine how any qualified applicant, regardless of physical or cognitive ability, can be effectively accommodated and counseled in achieving the most appropriate medical career. If carefully selected and supported, a student with a significant disability can succeed in a rigorous medical school program. Regardless of whether a person has a disability, each candidate for medical school must demonstrate that he or she has the potential to satisfy the key criteria of intelligence, professional attitude, and the ability to interact and communicate effectively, with or without reasonable accommodations and modifications.

Recommendation 1:
First and foremost, there is a need to reevaluate the goals and expectations of medical education and residency training to be consistent with the practice of medicine in the 21st century. Medical schools should modify the excessively strict technical standards that currently constitute a major barrier to many potential applicants. Indeed, one could argue that physical technical standards should not be required for graduation from medical school, but should be deferred to postgraduate education, where clear standards can be tied to the scope of practice of a particular specialty. The AAMC can help by updating its advisory materials with respect to the ADA experience over the past decade.

Recommendation 2:
More research is needed on people with disabilities in the health profession to determine the number of people with disabilities applying to medical school, and their rates of admission, graduation, and resultant professional experiences. Additional research is needed to identify the primary barriers to medical school and health professions for people with disabilities.

Recommendation 3:
A large, well controlled formal epidemiologic survey should be planned and implemented to accurately ascertain the prevalence of all degrees and types of physical disabilities among practicing physicians and medical students, as well as the effects of such disabilities on medical practice.\(^2^8\)

Recommendation 4:
Because of the dearth of medical literature, physicians and medical students with disabilities should be encouraged to document their own experiences and practice strategies to develop successful models. Such documentation would assist medical educators in constructing strategies for approaching reasonable accommodations and program modifications for medical students with disabilities. Such information could also have an effect on peer attitudes.\(^1^1\)

Recommendation 5:
Physicians and other healthcare professionals should be educated about the broad definition of disability, and encouraged to take steps to comply with the requirements of the ADA.

Recommendation 6:
Medical and other health professional schools should make a commitment to include and integrate clinical training and resources about disability throughout the educational process. The role and value of screening and preventive care for persons with disabilities needs to be emphasized.

Recommendation 7:
Medical and other health professional schools should integrate a disability curriculum into their medical training programs. With nearly one in every five people having a disability of some kind, it is imperative to raise the level of awareness and, consequently, the level of understanding about disability issues within the medical professions.

Recommendation 8:
Medical and other health professional schools should make meaningful efforts to promote the accessibility of their programs, sending the clear signal that people with disabilities are encouraged to apply—not just that they will be free from discrimination if they do.

Recommendation 9:
The AAMC should incorporate disability-related questions into the AAMC graduation questionnaire, to create benchmarks to measure improvement in this area.

Recommendation 10:
A task force(s) needs to be created to update standards and guidelines with respect to applicants with disabilities for admissions committees, licensure authorities, certifying boards, and privileging organizations. A study should be developed and implemented to determine whether the example of professionals (preferably physicians) with disabilities can effectively motivate patients with disabilities to set higher goals for rehabilitation, community integration, education, and employment.

CONCLUSION
Increasing inclusion of people with disabilities in the healthcare delivery system should not be viewed as an altruistic gesture, but as a matter of basic civil rights. Core principles of the ADA are nondiscriminatory inclusion and reasonable accommodation. For medical schools and training programs, this has profound implications—many
of which have not been fully recognized. Medical schools and other training programs are required to provide equal access to programs in the most integrated setting possible, accessible facilities and transportation, effective communication for teaching and training, and reasonable modifications to policies and procedures, including testing of students.

The ADA does not prevent medical schools from selectively accepting the most highly qualified applicants, nor does it impose any obligation on medical schools to lower their standards. The ADA does, however, protect applicants with disabilities from discrimination based on disability in the application process.39 Similarly, the ADA protects students with disabilities from being discriminated against as they matriculate. Finally, the ADA requires employers (such as hospitals or medical schools) to provide reasonable accommodations to their employees, including faculty and residents.29

In short, the ADA placed disability status on the same level as gender, race, and ethnicity in terms of federal nondiscrimination requirements. Just as it a moral charge to “take active steps to ensure that our healthcare practitioner community mirrors society’s gender, racial and ethnic mix,” it is equally imperative to extend this charge to people with disabilities. It is no less a matter of social justice and equality to incorporate and accommodate people with disabilities into the medical professions. Similar to race, gender, and ethnicity, incorporation of people with disabilities is a means to improve access to health care on the part of the underserved—people with disabilities. It is also a way to deliver “culturally” competent care—in that the disability community has developed a culture during the past several decades that mirrors those of other minority groups. And, finally, meaningful inclusion of people with disabilities in the medical professions just makes sense, and would fully employ the rich and diverse pool of our nation’s people to better manage the healthcare system.8,10

ACKNOWLEDGMENT

We would like to thank Kristi L. Kirschner, MD, for her input.

REFERENCES

2. Cohen JJ: Reconsidering disabled applicants. AAMC Reporter June 2004
24. Gonzales v National Board of Medical Examiners, 225 F3d 620 (2000) (Student with learning disability requested more time to take examination)
25. Ohio Civil Rights Commission v Case Western Reserve University, 666 NE2d 1376 (1996) admission of blind applicant
26. H.K. Wong v The Regents of the University of California, 192 F3d 807 (1999) (student challenged dismissal from medical school based on disability)
29. Stineman MG, Steinberg AG, Iezzoni LI, Conill A. The promotion of medical faculty with disabilities. AAP Newsletter, Spring 2003;7–10
Influence of the Prospective Payment System on Speech-Language Pathology Services

ABSTRACT


Objective: The present study was performed to determine the clinical effects of the Inpatient Rehabilitation Facility Prospective Payment System (IRF PPS) on speech and language intervention services and to examine the feasibility of using the federally mandated FIM™ instrument to establish resource allocation to patients with cognitive, communication, and swallowing disorders.

Design: A pre-IRF PPS and post-IRF PPS comparative study was conducted over a 1-yr time interval using data from the American Speech-Language-Hearing Association’s National Outcomes Measurement System. Toward this end, the National Outcomes Measurement System’s Functional Communication Measures were used to obtain data from 2,631 patients residing in 96 freestanding rehabilitation hospitals or hospitals with rehabilitation units implementing the prospective payment system on or after January 1, 2002. To ensure reliable retrospective and prospective data comparisons, all sites were active participants within the National Outcomes Measurement System program before the introduction of IRF PPS within their facilities.

Results: Findings revealed changes in both the utilization of speech-language pathologists and patient outcomes. Under the IRF PPS, there was a clear decline in speech- and language-related lengths of stay. However, clinicians attempted to compensate for these decrements in lengths of stay by increasing the intensity and frequency of their speech and language services. Despite these compensatory efforts, further analyses of the data revealed that under the IRF PPS, fewer patients achieved multiple levels of functional progress in speech and language abilities than before this payment system was implemented. This trend was most noteworthy in the treatment areas of swallowing, motor speech, and memory. In addition, this study revealed that, compared with the National Outcomes Measurement System’s Functional Communication Measures, the FIM instrument significantly under-represented and undervalued the extent of a patient’s overall progress in recovering from their cognitive, communication, or swallowing disabilities.

Conclusion: These findings support the notion that the introduction of the IRF PPS has, perhaps unintentionally, caused more patients with cognitive, communication, and swallowing disorders to be discharged from inpatient rehabilitative care with less than adequate functional skill levels. The discouraging results in speech-language pathology utilization and patient outcomes will be useful for clinicians in the future when facing the ongoing challenges of maintaining quality care while streamlining services under the prospective payment system.

Key Words: Inpatient Rehabilitation Facility Prospective Payment System, Speech-Language Pathology, Adult Patients, Rehabilitation
In 1983, Medicare, the single largest public healthcare program in the United States, introduced a national cost-containment measure known as the Prospective Payment System (PPS). The primary goal of the PPS was to reduce the ever-escalating growth of healthcare costs by establishing fixed payment rates for hospital care. For many hospitals, this meant eliminating the previously used, cost-based reimbursement approach and implementing a predetermined, flat-rate payment system in which providers received a set financial reimbursement for treating patients with a given illness. Initiated first in acute inpatient hospital settings, the inpatient PPS classified Medicare patients into diagnosis-related groups based on their primary medical diagnoses. The diagnosis-related groups then set the single rate of payment the hospital would receive for treating each patient, regardless of length of stay or type of care needed.

An early appraisal of the inpatient PPS program in acute hospital settings showed encouraging results. Under this controlled-payment system, healthcare providers were forced to streamline service delivery processes and thus became more accountable for resource utilization and spending. As a result of the inpatient PPS, evidence suggested a slowed growth in healthcare expenditures, along with a dramatic decline in hospital lengths of stays. At the same time, however, it seemed possible that these shortened lengths of stays could inadvertently give providers an increased incentive to minimize hospital costs, giving credence to what became known as the “quicker-and-sicker” theory. Many healthcare workers accepted the notion that shorter or “quicker” lengths of stay led to a reduction in ancillary and often necessary services, thus causing patients to be discharged in less stable or “sicker” medical conditions.

Initially, it was unknown if these speculations were valid—would the inpatient PPS create a ripple effect across the continuum of care? More specifically, would these initial cost savings be absorbed elsewhere in the healthcare system? Indeed, the outcomes of a number of subsequent studies indicated that the onset of the inpatient PPS in acute hospital settings triggered a surge in the use of postacute services. In fact, such results suggested that the initial cost-savings accrued from the inpatient PPS had, in fact, been absorbed by other levels of care with providers supplementing or replacing acute care treatment for treatment not yet subject to PPS. Although the initial economic outlook of the inpatient PPS seemed promising, overall aggregate spending across the continuum of care remained stable at pre-PPS levels. Thus, by the year 2000, healthcare costs continued to escalate at a dramatic rate, with ~13% of the gross domestic product being devoted to healthcare expenditures, with overall costs reaching 1.3 trillion dollars.

In an effort to further rein in the growth of healthcare spending and control spiraling costs, the United States Congress turned its attention to postacute treatment settings for the next wave of cost-containment. Early in August of 2001, the Centers for Medicare and Medicaid Services, formerly known as the Health Care Financing Administration, published the “Final Rule,” which set forth what is now known as the Inpatient Rehabilitation Facility Prospective Payment System (IRF PPS). The IRF PPS was launched on January 1, 2002, in the initially exempt freestanding rehabilitation hospitals and hospital rehabilitation units. In contrast to the diagnosis-related group–based system in which emphasis was placed on the patient’s primary medical diagnosis, this new system linked rehabilitation payment to the restoration of function. The IRF PPS utilized a function-based payment method by creating the “diagnosis-related group equivalency” known as case-mix groups. To categorize patients into similar clinical groupings or case-mix groups, the widely known FIM™ instrument is used in conjunction with the patient’s impairment level, age, and co-morbidities. Like its acute hospital predecessor, the case-mix groups establish the single unit of payment that providers now receive for each patient’s total rehabilitative care, regardless of amount or type of treatment needed. The premise behind the use of a function-based classification system is that patients with greater functional impairment levels would in turn need greater rehabilitation resources and services and, therefore, be placed in higher-weighted case-mix groups with higher payments rates.

A review of the literature revealed various studies showcasing the importance of the use of a functional status measurement tool rather than primary medical diagnosis as the main predictor of payment for this level of care. In a study by Hosek et al., the authors warned against the sole reliance on medical diagnosis to adequately capture resource use for inpatient rehabilitative care and discovered that medical rehabilitation costs and resource needs were too variable among patients with the same diagnosis, therefore, making it difficult for inpatient rehabilitation hospitals to operate under a diagnosis-based system. Instead, Carter et al. and others pointed to the importance of rehabilitation and its focus on the restoration of function and stated that a case-mix measurement system must “measure the extent of the functional deficit being restored.” Although many studies agreed with the general principle behind the use of
a function-based classification system to determine payment for medical rehabilitation, the introduction of the IRF PPS brought about a host of concerns regarding its impact on patient care. That is, if the IRF PPS implementation resulted in reduced lengths of stay, as the PPS seemed to have done in the acute inpatient hospital setting, would rehabilitation providers be able to shorten lengths of stay without compromising patient care and functional outcomes?

The effects of the PPS in other treatment settings have shown mixed results thus far. Although the results of a number of studies suggested that there were no adverse affects on quality of care as a whole, other findings indicate a range of qualitative changes. For example, whereas Rogers et al. found access to care and mortality after acute inpatient hospitalization to be virtually unaffected under the inpatient PPS, these investigators and others went on to note that patients had an increased likelihood of being discharged home in unstable or in less than optimal functioning conditions. In addition, further studies suggested that earlier discharges from acute care hospitals under the inpatient PPS led to increased admissions to long-term care facilities. To date, research has primarily addressed the overall qualitative changes to patient care or access to care in acute hospital settings, rather than focusing on any specific effects on patient outcomes. Since the inception of the PPS across the healthcare continuum, few studies have reported changes in therapy utilization and none have been directed toward the long-term effects of the PPS on patient outcomes or changes in speech and language service delivery.

The present study represents an initial attempt to determine the clinical effects of the IRF PPS in one area of rehabilitation that includes speech-language pathology interventions. To date, the evidence concerning changes in speech and language services under acute and postacute PPS treatment settings has been purely anecdotal. Clinicians working across the healthcare continuum have reported a decrease in speech and language referrals and staffing cutbacks and greater productivity demands under the PPS. In addition, speech-language pathologists (SLPs) working in inpatient rehabilitation settings historically have voiced significant concerns regarding the use of the FIM instrument, often criticizing its lack of sensitivity in measuring cognitive, communication, and swallowing skills. Based on this anecdotal evidence, we measured the effect of the IRF PPS on the use of speech-language pathology services and, more importantly, on patient outcomes. Therefore, the purpose of the present study was 2-fold: (1) to determine the primary effects of the IRF PPS on speech and language services and how this effects a patient’s recovery process and (2) to assess the suitability of the FIM instrument as the sole outcomes measurement tool in determining resource allocation to patients with cognitive, communication, and swallowing disorders.

METHODS

To accomplish these goals, the study relied on data collected using the American Speech-Language-Hearing Association’s (ASHA’s) National Outcomes Measurement System (NOMS) for Adults in Healthcare Settings. The NOMS comprises three national data collection systems designed to measure speech and language outcomes for various patient/client populations served by SLPs. The three NOMS components include Adults in Health-care, Pre-Kindergarten Children, and Children in Kindergarten through Grade 12. The cornerstone of NOMS is the use of ASHA’s Functional Communication Measures (FCMs), a series of disorder-specific, 7-point rating scales. Specifically, each scale contains seven discrete gradations of change, ranging from the least functional (level 1) to a most functional (level 7) level. The scales were developed for and by SLPs and are intended to measure functional progress and benefits from speech and language intervention. As such, the FCMs encompass the skill areas most commonly addressed by SLPs. To ensure FCM-rater reliability, only NOMS-certified clinicians can administer the NOMS evaluations. These clinicians must participate in a training program developed by ASHA and pass a NOMS user registration test to demonstrate understanding of the FCMs and their practical application in the field.

Unlike the FIM instrument, the entire set of FCMs is not intended to be measured at admission and discharge for each patient. Rather, a clinician selects specific FCMs based on the patient’s actual cognitive-communication or swallowing impairments and then scores the patient on the selected FCMs at the initiation of treatment. When SLP services are discontinued, these same FCMs are again rated. It is assumed that the difference, if any, in an FCM score from admission to discharge reflects the amount of functional gain a patient makes after speech and language intervention. In addition to these outcome measures, NOMS requires SLPs to provide information on each patient’s demographics, diagnosis, treatment setting, and service delivery mode and to submit each case to a national database maintained by ASHA. These data, in turn, provide clinicians with national comparisons on which to benchmark their performance.
For IRF PPS comparisons, NOMS data were collected from the Adults in Healthcare component, the largest of the three NOMS systems. At the time of the study, this database contained >50,000 cases provided by 1,600 registered NOMS users. For the purposes of this study, 96 data-collection sites were recruited for participation in our study. To ensure reliable retrospective and prospective comparisons, all sites were active participants within the NOMS program before the implementation of the IRF PPS within their facility. Thus, each site had been designated as a freestanding rehabilitation hospital or hospital with a rehabilitation unit activating the IRF PPS on or after January 1, 2002.

To fully assess the effect of the IRF PPS, in addition to the standard NOMS data (i.e., the FCM scores, details about patient demographics, diagnosis, service delivery and amount, frequency and intensity of services), SLPs also reported three further types of information relevant to pre-IRF PPS and post-IRF PPS comparisons. These additional data included FIM scores at admission, FIM scores at discharge, and the patient’s discharge disposition.

As part of the control procedures of the present study, sites also reported data on non-Medicare patients receiving SLP-initiated services in their inpatient rehabilitation settings and similar results from patients in other treatment settings covered by Medicare insurers. The collection of these data assisted in determining whether changes, if any, in the outcomes of Medicare patients post-IRF PPS were specific to that patient group or whether they were part of broader changes affecting patients in other treatment settings or patients covered by payers other than Medicare. These combined data were collected on a total of 2,631 patients during a 1-yr time span so that the data set included 1,912 completed admission and discharge cases and 719 admission-only cases. In addition, data were collected on 820 non-Medicare patients receiving speech-language pathology services in inpatient rehabilitation settings and 10,412 Medicare patients receiving speech-language pathology treatment in non-IRF PPS settings. Data were analyzed statistically and illustrated using SPSS (vs.10.0) software (SPSS, Chicago, IL).

RESULTS

Before examining the relationship between the IRF PPS and speech- and language-treatment outcomes, a comparison of the database before and after the implementation of the IRF PPS was made. As noted in Table 1, the populations before and after implementation of the IRF PPS were essentially identical on the key characteristics of patient age, sex, and medical diagnosis. Thus, there were no significant differences in sex or age for the average patient, nor in the distribution of the primary diagnosis categories across patients, when the pre-IRF PPS vs. post-IRF PPS databases were compared.

SLP Utilization

A comparison of pre-NOMS and post-NOMS data revealed a dramatic decline in speech and language lengths of stay under the IRF PPS. On average, the average length of stay post-IRF PPS for speech and language services was reduced by approximately 7 days. The bar-graph histogram of Figure 1 clearly shows that during the pre-IRF PPS period (black bars), more patients received SLP services for >30 days than post-IRF PPS (gray bars) \((\chi^2 = 94.3, df = 1, P < 0.001)\). In fact, >25% of the pre-IRF PPS patients compared with <10% of the post-IRF PPS patients received treatment from an SLP for >30 days. In addition, with the onset of the IRF PPS, the majority of patients received speech-language pathology services for fewer than 20 days. Moreover, this dramatic reduction in lengths of stay for speech-language pathology services was not evident among non-Medicare patients in inpatient rehabilitation settings \((\chi^2 = 0.74, df = 1, P = 0.390)\) nor for Medicare

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Pre</th>
<th>Post</th>
<th>(\chi^2)</th>
<th>Significance (P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>% male</td>
<td>47.1%</td>
<td>50.1%</td>
<td>1.098</td>
<td>0.295</td>
</tr>
<tr>
<td>Age at admission, &lt;60 yrs</td>
<td>28.9%</td>
<td>30.5%</td>
<td>0.492</td>
<td>0.483</td>
</tr>
<tr>
<td>Age at admission, &gt;80 yrs</td>
<td>9.6%</td>
<td>9.0%</td>
<td>0.119</td>
<td>0.730</td>
</tr>
<tr>
<td>Primary diagnosis, cerebrovascular disease</td>
<td>63.6%</td>
<td>58.4%</td>
<td>3.467</td>
<td>0.063</td>
</tr>
<tr>
<td>Primary diagnosis, respiratory disease</td>
<td>3.1%</td>
<td>4.0%</td>
<td>0.629</td>
<td>0.428</td>
</tr>
<tr>
<td>Primary diagnosis, hemorrhage/injury</td>
<td>5.4%</td>
<td>5.8%</td>
<td>0.052</td>
<td>0.820</td>
</tr>
<tr>
<td>Primary diagnosis, central nervous system disease</td>
<td>3.2%</td>
<td>4.3%</td>
<td>0.869</td>
<td>0.351</td>
</tr>
<tr>
<td>Primary diagnosis, head injury</td>
<td>4.0%</td>
<td>6.1%</td>
<td>3.005</td>
<td>0.083</td>
</tr>
<tr>
<td>Primary diagnosis, neoplasm</td>
<td>1.9%</td>
<td>2.7%</td>
<td>0.489</td>
<td>0.484</td>
</tr>
</tbody>
</table>
patients in acute hospital settings. In fact, lengths of stay among Medicare patients in acute hospital settings saw a significant increase ($\chi^2 = 23.9$, $df = 1$, $P < 0.0001$).

It could easily be assumed that the shortened lengths of stay for speech-language pathology patients would also lead to a reduction in speech and language services under the IRF PPS. At first glance, however, this hypothesis did not seem to hold true in that a comparison of the NOMS data for the periods before and after IRF PPS implementation revealed no significant effects on the amount of speech and language services patients received. Consequently, as seen in Figure 2, on average, after the implementation of the IRF PPS, patients continued to receive approximately 10 hrs of speech-language pathology services compared with the 11 hrs of treatment received before the execution of the new payment system. Thus, overall, there were few differences in the total amount of speech-language pathology treatment time before or after implementation of the IRF PPS.

However, with respect to speech-language pathology utilization, what did change under the IRF PPS was the intensity and frequency of speech and language services. These data represented in Figure 3 indicate that SLP services were significantly altered under the IRF PPS. During the pre-IRF PPS (black bars) period, the majority (89%) of speech and language patients received five or fewer treatment sessions per week. Under post-IRF PPS (gray bars) conditions, however, the majority (77%) of speech and language patients received more than five treatment sessions per week. This shift toward conducting more treatment sessions by SLPs post-IRF PPS was significant ($\chi^2 = 741.44$, $df = 1$, $P < 0.001$). Such findings suggest that SLPs compensated for the decreased lengths of stay by providing more frequent therapy sessions over a more condensed period of time. Again, this result seems limited to patients under the IRF PPS, as no significant differences in the frequency and intensity of speech-language pathology services were observed among non-Medicare patients in inpatient rehabilitation settings ($\chi^2 = 0.279$, $df = 1$, $P = 0.597$), nor among Medicare patients in other treatment settings ($\chi^2 = 0.483$, $df = 1$, $P = 0.487$).

**Patient Outcomes**

To effectively offset any changes in speech and language outcomes under the IRF PPS, it seems then that SLPs increased the number of their treatment sessions. Moreover, the initial implications from these modifications in the frequency of clinical service delivery seemed to be positive. That is, as shown in Figure 4, treatment-related progress in communication and swallowing abilities, whether improving one or more levels, essentially remained unchanged under the IRF PPS ($\chi^2 = 0.453$, $df = 1$, $P = 0.501$). The sole exception was the Spoken Language Comprehension FCM, in which 80% of patients made progress post-PPS, compared with only 67% pre-PPS ($\chi^2 = 8.5333$, $df = 1$, $P = 0.004$). Consequently, as measured by the NOMS' FCMs, patients receiving speech and language treatment were found to be just as likely to make functional progress post-IRF PPS as compared with their counterparts, who received similar treatment during the pre-IRF PPS period.

However, a more extensive analysis of the NOMS database with respect to the number of levels of treatment-related improvement exhibited by patients on average revealed troubling results.
Specifically, when comparing the degree of improvement that speech and language patients made under the IRF PPS, the data revealed a marked decline in patient outcomes. That is, whereas pre-IRF PPS and post-IRF PPS findings as illustrated in Figure 4 support the fact that treated patients progressed at least one FCM level in their speech and language skills, further analysis revealed that fewer patients achieved multiple levels of functional improvement after implementation of the IRF PPS ($\chi^2 = 45.330, df = 1, P < 0.0001$). As shown in Figure 5, this lack of progress was particularly noteworthy in the areas of motor speech, swallowing, and memory when compared with improvements within the other FCM scales.

Thus, patients were leaving their inpatient rehabilitation settings before achieving the higher levels of functioning they would have achieved before implementation of the IRF PPS. In addition, many more of these patients were leaving during the post-IRF PPS period needing further speech and language interventions and, in addition, were more likely to be discharged to home or to skilled nursing facilities rather than to other settings as illustrated in Figure 6. Interestingly, for non-Medicare patients receiving clinical services from SLPs in similar inpatient rehabilitation settings, no significant changes were observed in the pattern of discharge destinations. The majority of these patients, as they
were before IRF PPS, were discharged to their home settings.

**Measures of Functional Status**

The present study also revealed some compelling differences between the FCMs and their FIM counterparts. A side-by-side comparison, for example, of the FCMs and the speech and language FIM items indicate several disparities between the ranges of disorders measured. Table 2 highlights a number of abilities for which the FCMs measured progress in performing a particular cognitive/communicative related skill but for which there was no corresponding FIM measure (noted as NA). Clearly, the FIM instrument lacks finite measures related to speech production, including motor speech, voice, and fluency skills and specific measures of attentiveness, which are strongly associated with cognitive function and ability to perform activities of daily living. The present findings, consequently, revealed that 43% of speech-language pathology patients made progress in one or more of the four FCM skills areas that were not recognized by the FIM instrument (i.e., motor speech, voice, fluency, and attention), which were not reportable under the IRF PPS.

Moreover, as shown in Table 3, which indicates the agreement in skill areas that correspond between the two scales, the FIM instrument seemed to be less sensitive with respect to measuring treatment-related improvements than were the FCMs. Fifty-three percent (53%) of FIM and FCM ratings agreed as to the presence or absence of functional progress, 2% showed progress on the FIM instrument that was not captured by the FCMs, and 45% showed progress on the FCMs that was not captured by the FIM instrument.

**DISCUSSION**

The present study examined the impact of the IRF PPS in one area of rehabilitation (i.e., treatment for speech-language pathology). To this end, pre-IRF PPS and post-IRF PPS changes in the utilization of SLPs were compared using the NOMS’ FCM scales developed by ASHA. In addition, these FCMs were compared and contrasted to the FIM scales to assess whether the federally mandated FIM instrument adequately captured treatment-related improvements in speech and language.
language skills under the IRF PPS. The overall findings from these comparisons paralleled the results established by a number of studies conducted in other treatment settings or for other professional healthcare services. For example, Yip et al.\textsuperscript{28} found similar changes in physical and occupational therapy utilization under the skilled nursing facility PPS (SNF PPS). These authors concluded that patients receiving such services experienced shorter lengths of stay post-SNF PPS. That is, on average, patients received five fewer physical therapy days than did pre-SNF PPS patients. Other investigators studying acute hospital settings also revealed similar changes in practice patterns across related fields of rehabilitation. Such studies found that, like SLPs, other rehabilitation professionals were compensating for the inpatient PPS by significantly altering the frequency and intensity of their services. For example, Gray\textsuperscript{29} reported an increase in the use of occupational therapy services under the inpatient PPS, whereas Wilke et al.\textsuperscript{30} reported that “to keep patient recovery levels constant, intensity of inpatient care [by physicians] was forced to increase.” In addition, in a study comparing the effects of the PPS on SNFs, Hutt et al.\textsuperscript{32} noted that although patients experienced an increase in occupational therapy services under the SNF PPS, this enhanced service provision led to limited improvements and, therefore, a decrease in the rate of discharges to the community.

Two additional studies completed in acute inpatient settings also found trends similar to those reported for SLPs. When comparing pre-inpatient and post-inpatient discharges for acute hospital patients, Evans et al.\textsuperscript{26} noted a significant increase in skilled nursing referrals (64%) and a decrease in home-health referrals after the implementation of the inpatient PPS. Similarly, Schmidt et al.\textsuperscript{33} uncovered a 5% decrease in the percentage of acute care patients discharged to their home settings and a 44% increase in those referred to long-term care facilities post-inpatient PPS. Thus, like others under the PPS, patients receiving speech-language pathology services in inpatient rehabilitation settings experienced reduced lengths of stay, for which SLPs and other rehabilitative specialists attempted to compensate for by increasing the frequency and intensity of the services they provided. In addition, SLPs similarly altered their discharge dispositions, assigning more patients to home health care or to SNFs.

### Table 2

#### Disorders assessed by Functional Communication Measures (FCMs) vs. FIM™

<table>
<thead>
<tr>
<th>FCM</th>
<th>FIM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motor speech</td>
<td>NA</td>
</tr>
<tr>
<td>Voice</td>
<td>NA</td>
</tr>
<tr>
<td>Fluency</td>
<td>NA</td>
</tr>
<tr>
<td>Swallowing</td>
<td>Eating</td>
</tr>
<tr>
<td>Spoken language</td>
<td>Comprehension (auditory)</td>
</tr>
<tr>
<td>Spoken language</td>
<td>Expression (vocal)</td>
</tr>
<tr>
<td>Writing</td>
<td>Expression (non-vocal)</td>
</tr>
<tr>
<td>Reading</td>
<td>Comprehension (visual)</td>
</tr>
<tr>
<td>Attention</td>
<td>NA</td>
</tr>
<tr>
<td>Memory</td>
<td>Memory</td>
</tr>
<tr>
<td>Pragmatics</td>
<td>Social interaction</td>
</tr>
<tr>
<td>Problem solving\textsuperscript{a}</td>
<td>Problem solving</td>
</tr>
</tbody>
</table>

\textsuperscript{a}Implemented November 2002.

### Table 3

#### Agreement between Functional Communication Measures (FCMs) and FIM™ scores on the same patient scored at the same times

<table>
<thead>
<tr>
<th>FCM</th>
<th>FIM</th>
<th>FCM—No Progress</th>
<th>FCM—Progress</th>
<th>FCM—Progress</th>
<th>FCM—Progress</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>Motor speech</td>
<td>NA</td>
<td>87</td>
<td>18.0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Voice</td>
<td>NA</td>
<td>29</td>
<td>20.9</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Fluency</td>
<td>NA</td>
<td>5</td>
<td>38.5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Swallowing</td>
<td>Eating</td>
<td>53</td>
<td>16.3</td>
<td>193</td>
<td>59.2</td>
</tr>
<tr>
<td>Spoken language</td>
<td>Auditory</td>
<td>55</td>
<td>20.1</td>
<td>197</td>
<td>71.9</td>
</tr>
<tr>
<td>Spoken language</td>
<td>Comprehension</td>
<td>41</td>
<td>19.2</td>
<td>146</td>
<td>68.2</td>
</tr>
<tr>
<td>Writing</td>
<td>Nonvocal expression</td>
<td>1</td>
<td>50.0</td>
<td>1</td>
<td>50.0</td>
</tr>
<tr>
<td>Reading</td>
<td>Visual comprehension</td>
<td>1</td>
<td>33.3</td>
<td>2</td>
<td>66.7</td>
</tr>
<tr>
<td>Attention</td>
<td>NA</td>
<td>121</td>
<td>20.5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Memory</td>
<td>Memory</td>
<td>90</td>
<td>21.0</td>
<td>267</td>
<td>62.2</td>
</tr>
<tr>
<td>Pragmatics</td>
<td>Social interaction</td>
<td>6</td>
<td>7.8</td>
<td>53</td>
<td>68.8</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>489</td>
<td>19.2</td>
<td>859</td>
<td>33.7</td>
</tr>
</tbody>
</table>
However, the present study went one step further and examined how the above changes in speech-language pathology practice patterns affected the overall quality of patient care. The results of this detailed analysis revealed that the introduction of the IRF PPS had a limiting effect on the amount of functional progress patients make in resolving their communication and swallowing difficulties. That is, comparative analyses of the FCM and FIM tools uncovered major differences in their abilities to measure treatment-related improvements in speech and language skills, leading to the conclusion that the FIM instrument significantly underrepresents and undervalues the extent of a patient’s overall progress in recovering from their cognitive, communication, or swallowing disabilities, in particular. As a result, the introduction of the IRF PPS has, perhaps unintentionally, resulted in more patients with much greater dysfunction leaving inpatient rehabilitative care still in need of continued speech-language pathology services, which they are forced to seek at other levels of the healthcare system.

Many clinicians in the field of communication disorders argue that the FIM instrument, which emphasizes physical activities of daily living rather than the cognitive-communication-related activities, will lead to the misallocation of resources, noting that not all FIM items have an expected correlation to cost. For example, Carter et al. determined that the FIM motor scores revealed a much stronger association between cost and function than did the FIM cognitive scores. As such, these findings raise further concerns that providers will not be compensated appropriately for the extensive treatment services needed by patients with cognitive, communication, and swallowing impairments. Indeed, Rondinelli et al. concluded that the use of functional predictors such as the FIM instrument, directed primarily at cost-containment, will likely overlook potentially important gains in other domains.

Although functional status seems to be the best predictor of resource utilization for inpatient rehabilitation settings, it is clear that continued research in this area is necessary. Although Hosek et al. found support for the relationship between charges billed for rehabilitative care and functional status, their study was limited because it only focused on a few cognitive-communication areas and not the full range of skills addressed by SLPs. Further studies are needed to ascertain the extent to which patients with communication, swallowing, and cognitive impairments make progress in all clinical areas and treatment settings and to track changes in costs across the continuum of care. It is noteworthy that this study found an increase in lengths of stay among Medicare patients in acute care hospital settings at the same time that inpatient rehabilitation lengths of stay were declining. Further investigation of this finding would be valuable in determining whether and to what extent this offsets any cost savings realized in the inpatient rehabilitation setting. Unfortunately, the conclusions related to length of stay are necessarily complicated by uncertainty over the relative contributions of therapeutic intervention and spontaneous, or “natural,” recovery over time independent of intervention. Finally, additional studies are needed to determine the overall functional status of patients at the end of treatment across the healthcare continuum.

The results from the present study have major clinical implications for the field of speech-language pathology and for the rehabilitation community as a whole. Although the primary goal of the IRF PPS is economic in nature, its overall success or failure in curtailing healthcare expenditures and its impact on patient care will not be fully determined for years to come. These findings and continued research will assist clinicians with the ongoing challenges they face under the current IRF PPS process. By benchmarking the sweeping changes in speech-language pathology utilization and patient outcomes, SLPs can identify effective strategies that streamline service delivery costs at the same time that a high quality of patient care is maintained.

ACKNOWLEDGMENTS

We thank Brenda Lonsbury-Martin, Janet Brown, Floyd Roye, Tracy Schooling, Patricia Gardner, and Beverly Wang for their assistance and the 96 rehabilitation facilities and the many clinicians who participated in this project.

REFERENCES


Hand-Held Dynamometry in Persons with Tetraplegia
Comparison of Make- Versus Break-Testing Techniques

ABSTRACT

Objective: To compare make and break techniques for hand-held dynamometry in persons with tetraplegia. We compared the interrater and intrarater reliability, the relative forces, and variability between examiner techniques.

Design: Two examiners with no previous hand-held dynamometry training performed hand-held dynamometry on the elbow flexors or extensors of 19 persons with upper limb weakness secondary to tetraplegia, using break and make techniques. Testing was performed in two sessions separated by 10 mins. Simultaneous recording from an electrogoniometer placed across the elbow was obtained for a subset of participants.

Results: Break and make techniques both showed intraclass correlation coefficients exceeding 0.9 for interrater and intrarater reliability. The maximum expected difference for 95% of repeated measurements was 3.5 kg, with hand-held dynamometry strength values that averaged 7–11 kg. Average break/make ratios ranged from 1.38 to 1.49. The electrogoniometric data showed that the two examiners used similar testing technique, and small variations in technique were not associated with significant differences in strength recordings.

Conclusion: Make and break techniques for hand-held dynamometry both show high reliability over a short intersession period when performed by inexperienced examiners on weak elbow flexors and extensors.

Key Words: Spinal Cord Injuries, Muscle Strength, Muscle Spasticity, Reproducibility of Results
Muscle strength testing is an important clinical technique for patient assessment and for research on treatment of diseases and injuries. Measurement of strength allows identification of muscle weakness, and muscle strength can be followed over time to monitor for improvement with treatment or decline with disease progression. Ideally, the method used for strength measurement should be easy to learn, portable, reliable, and sensitive to change. Manual muscle testing (MMT) meets some of these requirements; however, it requires considerable training to establish good interobserver consistency, and for some muscle groups, the reliability of testing has been questioned. Also, the 6-point (0–5) Medical Research Council scale is nonlinear and relatively insensitive to change, as grade 4 strength may be assigned with as little as 10% of predicted normal strength. Quantitative strength measurement using a force gauge, either fixed or as a component of an isokinetic dynamometer, is currently considered the criterion standard for strength measurement. Disadvantages of these techniques include difficulty measuring strength in some muscles and the cost and lack of portability of isokinetic testing equipment.

Hand-held dynamometry (HHD), also known as myometry, combines some advantages of both MMT and conventional force-gauge testing. It is portable and can be used to test multiple muscles. Previous studies have generally shown high intrarater and interrater reliability when tested in various patient populations and in normal participants. It is also superior to MMT for detection of mild and moderate weakness and changes in muscle strength. Testing is performed using one of two techniques, termed “make” and “break.” Make technique requires the examiner to resist a maximal voluntary contraction by the examinee and is essentially an isometric contraction. With break technique, the examiner applies adequate force to overcome the examinee, thereby producing an eccentric contraction. Because eccentric strength is greater than isometric, break technique produces somewhat greater strength values. A number of investigators have assessed the reliability of the two techniques, and generally, both have shown similar reliability. Despite this, there may be other factors that make one technique preferable. For example, a weak examiner performing make-technique testing may not provide enough resistance to allow the subject to exert a maximal contraction, resulting in an erroneously low strength measurement; if examiner strength is inadequate for break-technique testing, the examiner will be unable to overpower the test muscle, and the examiner becomes aware that the test is underestimating muscle strength.

A number of questions remain regarding HHD and the two testing techniques. First, although most research shows excellent reliability, nearly all have used highly or moderately experienced examiners, and none have compared make and break technique in novice examiners. A number of previous studies have addressed reliability of the individual techniques in examiners without experience performing HHD. One showed high test–retest reliability for novice examiners using make technique after completion of a 3-hr training session. Intraclass correlation coefficients (ICC) of 0.83–0.86 have been reported for a single examiner using break technique after <5 hrs experience with HHD. An additional study demonstrated good reproducibility for a novice examiner who performed HHD against a spring-loaded device that would break at specified forces. We are unaware of any studies that have examined the intrarater reliability of break-technique HHD on patients when performed by inexperienced examiners.

A second issue related to HHD that remains unresolved is the difference in strength recordings with break vs. make technique. Previous studies are not in agreement, with the break exceeding make by <10% in some studies or as much as 70% in others. An additional factor not previously assessed is the degree to which differences in performance of break technique could affect the break/make ratio. If the examinee is overcome during break technique at a higher speed (angular velocity), this may cause a greater strength value to be recorded than if done at a lower angular velocity. Previous studies provide only brief descriptions of how break technique was performed, with regard to how fast the examinee was overcome, and descriptions of break technique vary considerably between studies.

The objective of this study was to compare the two HHD techniques (make and break) when performed by inexperienced examiners on persons with tetraplegia. We assessed the intrarater and interrater reliability of the two techniques, the relative forces recorded, and technique-related factors that could affect the recorded strength values and the reliability of the measurements. We speculated that make technique could show better reliability than break technique, due to greater variability in how novice examiners perform testing using the break technique.

**METHODS**

**Participants**

Study participants were recruited from the inpatient spinal cord injury ward at the participating...
institution. Inclusion criteria included a history of spinal cord injury of traumatic or nontraumatic pathogenesis and weakness of either the elbow flexors or extensors with MMT grades of 3/5 or 4/5. Exclusion criteria included upper limb pathology that would preclude maximal strength testing (e.g., acute fracture), cognitive dysfunction with inability to reliably follow commands, or medical contraindications to maximal strength testing. During the data collection period, all inpatients on the ward who met study inclusion and exclusion criteria were offered participation. The study protocol received approval from institutional review boards at the participating institution. Written informed consent was obtained from all participants.

Each participant underwent clinical examination of bilateral upper limbs by the senior author for strength, reflexes, and spasticity. MMT was performed for the American Spinal Injury Association (ASIA) key muscles, including elbow flexion and extension, using whole-number muscle grades as described in the ASIA reference manual. Participants were classified for motor level and ASIA Impairment Scale scores. For participants with asymmetric motor level, the more rostral motor level was recorded. Deep tendon reflexes (DTRs) for bilateral biceps and triceps were assessed using a tendon reflex hammer and graded on a 0–3 scale (absent, hypoactive, normal, hyperactive). Spasticity was assessed for elbow flexion and extension by passively moving the upper limb and grading the degree of resistance to movement using the modified Ashworth scale (MAS). Based on the clinical examination, we selected a single muscle group (either elbow flexion or extension) on the right or left side for subsequent testing. As we intended to determine the association between upper motor neuron signs and strength measurements, and we have observed that these signs are relatively infrequent near the zone of injury in persons with spinal cord injury, we preferentially selected muscles with elevated MAS or DTR scores. Muscle selection was limited to those with MMT scores of 3/5 or 4/5. Of these muscles, we selected the one with the greatest abnormal (>0) MAS score, and in cases in which two or more muscles had the same abnormal MAS score, we chose the weakest muscle. If all weak muscles had a MAS score of 0, then we chose the muscle with greatest DTR score, similarly choosing the weakest muscle in cases of equal DTR scores.

Instrumentation and Testing Protocol

During testing, the participants were positioned supine on a hospital bed, with the shoulder abducted 20 degrees, the posterior arm contacting the bed, the forearm in vertical orientation, and a neutral position of pronation-supination. A TSD 130B electrogoniometer (BIOPAC Systems, Goleta, CA) was placed across the elbow joint, and a pre-amplified surface electromyogram electrode (TSD 150, BIOPAC) was placed over the muscle group undergoing testing (Fig. 1). Electrogoniometer calibration was performed with the elbow at 90 degrees and at maximal extension, as determined with a universal goniometer aligned with the humerus and radius, with the fulcrum centered over the lateral epicondyle. The electromyogram was used to confirm periods of participant muscle activity but was not otherwise analyzed. HHD was performed for either elbow flexion or extension using a Chatillon CSD-200 dynamometer, (AMETEK, Paoli, PA) with the distal edge of the HHD pad aligned with the radial or ulnar aspect of the radiocarpal joint. The peak force shown on the digital display of the HHD was recorded. During the strength-testing procedure, the analog outputs of the HHD, electrogoniometer, and electromyogram electrode were sent to a MP100WS interface, sampled at 500 Hz, and recorded using AcqKnowledge software (BIOPAC). Force was measured in Newtons, and it is reported in terms of the gravitational force acting on a specified mass (in kilograms).

Two physical therapy students each performed testing on each participant. Previous examiner experience with HHD involved testing one muscle on a normal individual. The testing protocol is shown in Figure 2. During the first part of the testing session, the first examiner tested the participant four times, twice using the make technique, and twice using the break technique, with a 10-sec rest period between each strength test. After a 30-sec rest, the second examiner performed testing using the same techniques. This procedure was repeated by the two examiners after a 10-min rest period. The order with which the examiners performed the testing was alternated for successive participants. The entire testing session took approximately 30
mins per participant. Consistent verbal encouragement was provided to participants by the examiners during testing. Both examiners were present throughout all testing.

Analysis

Goniometric data were reviewed graphically along with the analog HHD output and electromyographic data. For each trial, we identified periods of peak force application on the HHD signal tracing and confirmed this with review of the electromyographic data. For break-technique trials, we calculated the angular velocity for the portion of the peak force application during which the participant’s muscle was overcome into an eccentric contraction, and we recorded initial and final joint angles for this movement. For make-technique trials, we determined whether the examiner had in fact performed a break-technique trial with excessive eccentric movement, defined as ≥5 degrees during the period of peak force application. If this occurred, we calculated angular velocity, initial angle, and final angle as with the break-technique trials.

In keeping with our clinical practice and that of other investigators, we report the analysis using the highest strength measurements from each trial. Previous investigations have compared use of the maximum, mean, or median for repeated HHD trials, or have compared the mean with a single trial, and have shown similar reproducibility with each method. Our measures of repeatability were only slightly lower if we restricted the analysis to the second of each pair of strength measurements. We chose to repeat contractions only twice per trial as we wished to minimize muscle fatigue. As there was no significant difference in strength of elbow flexors vs. extensors (P > 0.30 for all comparisons), we chose to combine data from both muscle groups for the analysis. Because electrogoniometer data were incomplete for many participants, we restricted the analysis of these data to participants who had complete data for both testing sessions, and we analyzed the goniometric data from each of these trials to characterize examiner technique. Differences in strength measurements were compared using t tests (two-tailed, P < 0.05 as significant), although similar results were obtained when assessed using non-parametric statistical tests. Reliability between measures was assessed with the ICC two-way mixed-effects model for intrarater assessments and the two-way random-effects model for interrater assessments. We calculated within-participant standard deviation to compute repeatability for pairs of measurements. We also plotted differences vs. mean values for strength values, as described by Bland and Altman, to assess for systematic differences across the range of strength values. Spearman’s rank-correlation coefficient was used to assess associations between spasticity and relative forces with the two techniques and between differences in angular velocity during break technique and relative strength. All statistical analyses were performed using SPSS 10.0.5 (SPSS, Chicago, IL).

RESULTS

Participants

A total of 19 individuals participated in the study. The mean participant age was 53.5 ± 11.7 yrs, and all participants were men. Of the 19 participants, three were undergoing initial spinal cord injury rehabilitation and were <6 mos postinjury, and the others were >1 yr postinjury. The motor level was C4, C5, and C6 for 1, 12, and 6 participants, and ASIA Impairment Scale scores were A, B, and D for 6, 3, and 10 participants, respectively. Table 1 shows the examination findings for the tested muscles. Although increased DTRs were prevalent in participants, high spasticity in tested muscles, as measured by MAS, was uncommon.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Elbow Flexors (n = 7)</th>
<th>Elbow Extensors (n = 12)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMT 3/5</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>MMT 4/5</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>DTR 0</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>DTR 1+</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>DTR 2+</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>DTR 3+</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Spasticity (MAS)</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>Spasticity (MAS)</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Spasticity (MAS)</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Spasticity (MAS)</td>
<td>3</td>
<td>0</td>
</tr>
</tbody>
</table>

MMT, manual muscle test; DTR, deep tendon reflex; MAS, modified Ashworth scale.
Interrater and Intrarater Reliability

Table 2 lists strength recorded for each muscle group, testing technique, session, and examiner. For both make and break techniques, strength measurements showed high reliability for both intrarater and intrarater reliability comparisons (Table 3). Representative data, showing the linear correlation between measurements for examiner 1 vs. 2 using break technique during session 1, are shown in Figure 3. Although the values were highly correlated, examiner 2 tended to record higher strength than examiner 1, with a significantly greater strength determined for both make and break techniques in session 2 in our combined analysis of elbow flexor and extensor data. During that session, mean ± standard deviation strength for make technique for examiners 1 and 2 were 6.9 ± 3.4 and 7.8 ± 3.8 kg and, for break technique, were 9.9 ± 5.1 and 10.9 ± 4.7 kg, respectively (P = 0.001 and 0.012 for comparisons). The analyses of technique factors that could have contributed to these differences are described below under “Electrogoniometer Data.” Bland–Altman plots of differences vs. mean values for strength measurements showed no systematic differences over the range of strength values. Data for break-technique testing from session 1 are shown in Figure 4.

We performed additional analyses to determine the repeatability of the measurement and to express it in a more clinically relevant way. For these intrarater comparisons, the mean difference in strength between the two sessions averaged between 1.0 and 1.5 kg. Within-participant standard deviation for the intrarater comparisons ranged from 0.8 to 1.3 kg, with no significant difference between make and break technique. The corresponding range for repeatability, 2.6–2.9 kg, indicates the maximum expected difference between two repeated measurements for 95% of paired observations. Within-participant standard deviation for the interrater comparisons ranged from 0.9 to 1.1 kg, with repeatability of between 2.6 and 2.9 kg, and there was no significant differences between repeatability for make and break techniques. Plots of differences between strength measurements for intrarater and interrater comparisons showed no evidence of improved agreement later in the study, after the examiners gained experience with HHD.

Break/Make Ratios

The break technique produced significantly greater strength measurements than did the make technique. We calculated this difference as the break/make (B/M) ratio. B/M ratios showed considerable variability between participants, and there was no significant difference in mean B/M ratios determined by the two examiners. For examiner 1, mean ± standard deviation B/M was 1.41 ± 0.39 in session 1 and 1.48 ± 0.38 in session 2, with corresponding values of 1.38 ± 0.29 and 1.49 ± 0.37 for examiner 2. We found no association between the B/M ratio and either the DTR or MAS for the test muscle.

Examiner Testing Technique: Electrogoniometer Data

We assessed electrogoniometer data to characterize and compare testing techniques between the two examiners and to determine whether differences explained the somewhat greater strength values obtained by examiner 2. For many participants, the electrogoniometer did not maintain calibration for an entire testing session, so these results are restricted to the eight participants with complete goniometric data for all trials through both sessions. This analysis includes all make- and break-technique measurements during those trials, involving 64 individual trials per examiner.

During make technique, both examiners had a similar likelihood of unintentionally moving the limb eccentrically through periods of peak force application, thereby convert-

<table>
<thead>
<tr>
<th>TABLE 2</th>
<th>Mean ± standard deviation muscle strength (in kilograms) by examiner, session, technique, and muscle group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Examiner 1</td>
</tr>
<tr>
<td>Session 1</td>
<td></td>
</tr>
<tr>
<td>Elbow flexors: make</td>
<td>7.5 ± 4.0</td>
</tr>
<tr>
<td>Elbow flexors: break</td>
<td>9.2 ± 4.8</td>
</tr>
<tr>
<td>Elbow extensors: make</td>
<td>7.6 ± 3.6</td>
</tr>
<tr>
<td>Elbow extensors: break</td>
<td>10.8 ± 5.2</td>
</tr>
<tr>
<td>Session 2</td>
<td></td>
</tr>
<tr>
<td>Elbow flexors: make</td>
<td>6.7 ± 3.9</td>
</tr>
<tr>
<td>Elbow flexors: break</td>
<td>8.7 ± 4.0</td>
</tr>
<tr>
<td>Elbow extensors: make</td>
<td>7.0 ± 3.3</td>
</tr>
<tr>
<td>Elbow extensors: break</td>
<td>10.6 ± 5.7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE 3</th>
<th>Reliability of strength measurements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intrarater reliability</td>
<td>ICC</td>
</tr>
<tr>
<td>Examiner 1 make</td>
<td>0.91</td>
</tr>
<tr>
<td>Examiner 2 make</td>
<td>0.94</td>
</tr>
<tr>
<td>Examiner 1 break</td>
<td>0.94</td>
</tr>
<tr>
<td>Examiner 2 break</td>
<td>0.93</td>
</tr>
<tr>
<td>Interrater reliability</td>
<td></td>
</tr>
<tr>
<td>Session 1 make</td>
<td>0.94</td>
</tr>
<tr>
<td>Session 2 make</td>
<td>0.97</td>
</tr>
<tr>
<td>Session 1 break</td>
<td>0.95</td>
</tr>
<tr>
<td>Session 2 break</td>
<td>0.94</td>
</tr>
</tbody>
</table>

ICC, intraclass correlation coefficient.
ing the trial into a break-technique trial. This occurred in 25% of trials for each examiner, and the mean angular velocity during peak force for these unintentional break trials was 18.7 degrees/sec for examiner 1 and 14.8 degrees/sec for examiner 2 ($P = 0.37$). We next analyzed the angular velocity at which the examiners performed break-technique trials. We found no significant difference in mean angular velocity between the two examiners ($37.4 \pm 15.7$ vs. $27.5 \pm 39.8$ degrees/sec for examiners 1 and 2, respectively; $P = 0.16$), although examiner 2 tended to show greater variability for angular velocity. In addition, when pairs of break-technique trials by the same examiner were compared, we found no association between the difference in angular velocity and the difference in strength measured. Finally, we saw no difference in the total degrees through which the limb was moved during peak force application for break technique (30.8 ± 18.2 vs. 31.3 ± 18.1 degrees for examiners 1 vs. 2; $P = 0.92$).

**DISCUSSION**

Although our study used novice examiners, we showed excellent interrater and intrarater reliability for HHD, as assessed with the ICC, when performed using both make and break techniques. These findings are in agreement with most previous
studies that have assessed experienced examiners or have provided considerably more training to examiners. It is possible that repeatability would be improved with additional examiner experience, but a trend toward better repeatability was not evident over the course of testing 19 participants. Others have shown an examiner with 2 yrs of experience performing HHD had better repeatability than less-experienced examiners, although testing in that study was performed against a fixed object. The current study is the first to compare reliability for the two techniques when performed by inexperienced examiners. Because make and break techniques had similar reliability, our results indicate no clear reason to choose one method over the other. Other investigators have noted examiner strength to affect the reliability of strength measurements on stronger muscles when make technique is used. Because we tested muscles that were relatively weak, we were unable to determine whether the two techniques would have similar reliability on strong muscles. Although the mean difference between repeated measurements was relatively small (1–1.5 kg), larger differences occasionally occurred. Using methods recommended by Bland and Altman, we determined that 95% of paired repeated measurements would vary by <3.5 kg. Therefore, a change of greater than this magnitude is more likely to represent a true increase or decrease in strength, provided that the patient is similar in characteristics to the population we studied, which had mean strength of approximately 7–11 kg. We think that expressing repeatability in these terms, rather than solely as an ICC, has greater clinical applicability for practitioners. The large magnitude of the repeatability measure in relation to the measured strength (3.5 kg for muscles with 7–11 kg of strength) indicates that relatively large differences in strength will occasionally occur on repeated measurements. This contrasts with the high ICC obtained for these measurements, and this highlights a limitation of the ICC as an indicator of measurement reliability.

Previous work has shown that the break technique produces greater strength values than the make technique, but there has been lack of agreement on the ratio between break and make values. Two studies have reported B/M ratios of <1.1, whereas others have reported ratios of around 1.3. Our findings are similar to the latter studies, with B/M ratios ranging from 1.38 to 1.49. Although we found no correlation between muscle spasticity and B/M ratios, our participants generally had low Ashworth scores for the muscles tested. We also found no relationship between DTRs for the test muscle and the B/M ratio. Because our sample size was small and there was a low prevalence of spasticity, this study had inadequate power to rule out a correlation between the B/M ratio and spasticity. Bohannon has reported considerably higher ratios (1.7) in paretic muscles in patients with stroke, and the ratios were positively correlated with Ashworth scores for muscle spasticity. If such a relationship is in fact present in patients with spinal cord injury with more severe spasticity, it would have potential clinical relevance. Modulation of spasticity with medications could have a differential effect on strength as measured with the two testing techniques. Although we determined mean B/M ratios, the two testing techniques should not be used interchangeably by applying this ratio as a conversion factor because the ratio varied considerably among individuals. Repeated measurements should be done using the identical technique, either break or make, to allow accurate comparisons, and documentation of HHD results should always specify the technique used.

We are unaware of previous studies that have quantified the movement that occurs during break-technique testing. Descriptive information on how break technique has been performed in those studies has been limited and we hypothesized that differences in technique could explain some of the variability in B/M ratios previously reported. In particular, we anticipated that break-technique testing performed at a higher angular velocity would result in greater strength measurement and thus a higher B/M ratio. Our two examiners performed break techniques in a similar manner according to the quantitative data we collected with the electrogoniometer. Mean angular velocity achieved during break technique differed by about 10 degrees between the two examiners (37.4 vs. 27.5 degrees/sec), and each examiner moved the tested limb through about 31 degrees during the period of peak force application. In addition, during make technique, both examiners had similar likelihood of unintentional movement, albeit of lower angular velocity and excursion than with intentional break technique. For the participants who had full goniometric data, we saw no evidence that either greater angular velocity during break technique or unintentional eccentric movement during make-technique testing were associated with higher strength recordings. The small number of subjects with goniometric data limits the ability to detect an association between strength measurement and angular velocity. It is possible that significantly higher or lower angular velocities could affect strength recordings, but within the range used by these two examiners, it was not found to be a factor. An additional limitation to our study is that the examiners viewed each other’s testing technique, and this could have resulted in greater technique consistency between the two examiners.
There are a number of additional limitations to our study. Participants were drawn from a patient population consisting of male veterans with primarily nonacute spinal cord injury. In addition, the examiners included only two female physical therapy students. The nonrandomized order in which techniques were used (make followed by break) may have biased the results, and it would be preferable to have randomized the order of techniques. Examiners were not blinded to the results of the other examiner or to their results from previous trials. Because the strength value we recorded was the peak force digitally displayed by the HHD, we do not think that this would be biased by the examiners’ knowledge of earlier strength recordings; others have drawn similar conclusions about the likelihood of this bias. However, as noted above, the presence of both examiners may have led to use of a similar testing technique. The time interval between sessions was short, and a greater interval could have resulted in lower reliability for the measurements. The repeated strength tests on weak muscles over a short time period may have caused muscle fatigue. Due to the small sample size, we combined data from elbow flexors and extensors for our analysis; reliability and B/M ratios could differ between these two muscle groups. Some previous studies have shown HHD reliability to be worse for lower limb muscles in normal subjects, but our study was restricted to two weak upper limb muscles, the elbow flexors and extensors. These muscle groups are among the easiest to test using HHD, so these findings cannot be generalized to all muscle groups.

CONCLUSION
Make and break testing techniques both show high intrarater and interrater reliability when performed by inexperienced examiners. Due to the minimal training needed, combined with the previously recognized advantages for detection of mild weakness and small strength changes, HHD should be considered as an alternative to MMT when strength testing is performed by less experienced examiners.

REFERENCES
Unintentional Vascular Uptake in Fluoroscopically Guided, Contrast-Confirmed Spinal Injections
A 1-Yr Clinical Experience and Discussion of Findings

ABSTRACT

Objective: Documentation of vascular uptake on spinal injection in the context of negative aspiration and negative passive filling of blood into the hub of the needle.

Design: A total of 1,295 consecutive outpatients receiving fluoroscopically guided, contrast-confirmed injection in a multispecialty practice over a 1-yr time frame were retrospectively reviewed with passive observation for inadvertent vascular uptake, passive filling, and required repositioning.

Results: Positive vascular uptake was seen in 2–13% of cases with variable degrees of aspiration, passive filling, and required needle repositionings to avoid vascular uptake.

Conclusion: Negative aspiration and allotment for passive filling is inadequate to confirm the absence of vascular injection. Spinal injection will never be risk free. The safest method is fluoroscopically guided, contrast-confirmed injection, which should be considered the current standard of care.

Key Words: Epidural Steroids, Radiculopathy, Spinal Injection, Pain Management, Fluoroscopy, Complications, Epiduragram
The first report of therapeutic spinal injection in humans, published in 1901, was accomplished without image guidance. Advances in radiology and pharmacology have propelled us quite a distance from last century’s “blind” spinal blocks. The use of real-time imaging and contrast agents to assist with needle placement and confirmation of flow in selective spinal injection has become common. Intravascular placement of multi-orifice epidural catheters with a negative aspiration and positive epinephrine response has been studied in women in labor, as has the accuracy of epinephrine test doses in the face of negative blood aspiration. What has not been studied is noting vascular uptake with contrast agent and repositionings in the context of negative aspiration and allotment of time for passive blood flow into the hub of the needle.

METHODS
All procedural reports, in a private multispecialty practice, from a consecutive series of all outpatients undergoing fluoroscopically guided, contrast-confirmed injections, over a 1-yr period at two sites, were retroactively reviewed by the primary author. Data detailing the 1,295 injections are provided in Table 1. All injections were performed by the primary author. The aspiration noted in this study was performed after placement by anatomic localization under fluoroscopic guidance or, in the case of translaminar epidurals, air loss-of-resistance technique in conjunction with fluoroscopic guidance. In accordance with the primary author’s standard practice, when a vascular pattern was observed under real-time injection of contrast agent, gentle aspiration was attempted and the results noted, the connection tubing was removed, and a 10-sec observation was made for passive filling into the clear plastic hub of the needle. After this, gentle re-aspiration was attempted. The needle was then repositioned to accomplish the planned injection without vascular uptake. The number of repositionings was recorded. Simple rotation along the axis of the needle was not considered to constitute a repositioning. Loss of resistance in translaminar epidurals demonstrating nonpuderal contrast patterns were disregarded, regardless of vascular uptake. No accidental subdural/subarachnoid injections of steroid/anesthetic were performed during the course of the study. One patient with known chronic airway disease developed chest pain 30 mins after cervical translaminar injection and was transferred to an emergency room. One patient after cervical facet injection became nauseated and required intramuscular promethazine.

Instrumentation and Equipment
All procedures were performed at an outpatient surgery center or community-based hospital. Either an OEC 9800 C-arm imager (Salt Lake City, UT) with 41% iopamidol injection (Braco Diagnostics, Princeton, NJ) or a Shimatsu MH100 free-standing angiographic system (Tokyo, Japan) with iothalamate meglumine U.S.P. 60% (Mallinckrodt, St. Louis, MO) were used for all procedures. Three and one-half–inch 20-gauge Weiss Touhy needles were used for all translaminar (cervical and lumbar) injections. Five-inch 22-gauge spinal needles were used for lumbar sympathetic block. Three and one-half–inch 25-gauge spinal needles were used for cervical facet injections and 3.5- or 5-inch (depending on body habitus). Three and one-half–inch 22-gauge spinal needles were used for all other injections. A Perifix Epidural Catheter (B. TABLE 1 Injection data

<table>
<thead>
<tr>
<th>Injection Type</th>
<th>n</th>
<th>n Vasc</th>
<th>Tn Vasc</th>
<th>% Vasc</th>
<th>Multiple</th>
<th>Aspiration</th>
<th>Passive Filling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lumbar translaminar</td>
<td>468</td>
<td>14</td>
<td>18^a</td>
<td>3.0</td>
<td>4</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>Lumbar transformalinal</td>
<td>178</td>
<td>16</td>
<td>22</td>
<td>8.9</td>
<td>6</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Lumbar Facet</td>
<td>150</td>
<td>13</td>
<td>13</td>
<td>8.7</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Sacral joint</td>
<td>177</td>
<td>12 (13)^b</td>
<td>15 (24)^b</td>
<td>7.3</td>
<td>5 (6)^b</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Cervical facet</td>
<td>108</td>
<td>10</td>
<td>11</td>
<td>9.3</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Cervical transaminar</td>
<td>100</td>
<td>2</td>
<td>4</td>
<td>2.0</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Lumbar sympathetic</td>
<td>100</td>
<td>2</td>
<td>4</td>
<td>2.0</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Intercostal nerve block</td>
<td>56</td>
<td>3</td>
<td>3</td>
<td>5.4</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Caudal with catheter</td>
<td>45</td>
<td>5 (6)</td>
<td>7 (12)</td>
<td>13.3</td>
<td>4</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

n Vasc, number of patients with at least one vascular pattern; Tn Vasc, number of vascular patterns; % Vasc, percentage of patients with vascular patterns; Multiple, number of times more than one repositioning was required (subset of n Vasc); Aspiration, number of patients with at least one vascular pattern positive for blood aspiration; Passive Filling, number of patients with at least one vascular pattern with passive filling of blood into the hub of the needle.

^aTwo (not included in Tn Vasc = 18) pulsatile arteriovenous malformation patterns on epidurogram.

^bOne occasion aborted after nine repositionings with continued vascular uptake.
Braun Medical, Bethlehem, PA), inserted through a standard 16-gauge angiocatheter, was used for caudal with catheter (CEC) procedures.

RESULTS

There were 450 patients aged 19–96 yrs, with an average age of 57 ± 16 yrs. A total of 109 vascular patterns were seen in 78 patients, with an average age of 65 ± 16 yrs, during a total of 1,295 injections (Table 1). Eight of the 109 patterns had positive blood aspiration, whereas 24 had passive flow of blood into the plastic needle hub. No positive re-aspirations were observed after the delay for passive fill, when the initial aspiration was negative. The greatest percentage of vascular patterns was seen with CEC, which also had the highest percentage of repeat vascular uptake, with four of six cases with vascular uptake requiring multiple (two to five) repositionings. The lowest percentage of vascular patterns was seen in cervical translaminar epidurals. The greatest number of vascular patterns in a single procedure was for an sacroiliac joint, requiring nine repositionings, none of which had a positive aspiration or passive fill, at which point the procedure was aborted.

On three occasions, when using a lumbar translaminar approach, what appeared as a pulsatile arteriovenous malformation on epidurogram necessitated CEC procedures. Two of these were without intravascular uptake; one had a mixed vascular/nonvascular epidural pattern. One of the two cases without vascular uptake on the initial lumbar translaminar approach had repeated vascular uptake into the low lumbar radiculomedullary veins with CEC and required three repositionings. The authors are uncertain of the clinical significance of this finding.

Of the 13 total vascular patterns on the lumbar facet, six had vascular patterns either at the adjacent facet or sacroiliac joint. One of the four vascular patterns with passive filling of the hub on lumbar facet was arterial.

On one occasion, from a C7/T1 right paramedian cervical translaminar approach, a pulsatile arterial pattern crossed the midline and extended rostrally to C3. Three translaminar passes were required to avoid this particular vascular pattern. Four of the eleven vascular patterns seen during cervical facet injections were at the level of C2/3. Only one required multiple positioning. All four cervical facet injection patterns were nonpulsatile and extraspinal. None had positive aspiration or passive filling.

One pattern during an intercostal nerve block traveled laterally, then sharply cephalad, to highlight the superior adjacent intercostal vessel. The

![FIGURE 1 Right L5 transforaminal injection with pulsatile arterial pattern evident on real-time fluoroscopy.](image_url)
other two patterns seemed to be low-pressure venous systems and traveled medial to the lateral aspect of the vertebral body.

**DISCUSSION**

The spine and anatomically affiliated structures have a highly variable blood supply. Spinal segmental arteries come off the aorta and form intercostal, ascending cervical, and lumbar arteries. They then give rise to the radiculomedullary arteries, which make a characteristic hairpin turn as they contribute to both the anterior and, to a lesser degree, posterior spinal arteries. The largest of these is the artery of Adamkiewicz, usually arising from a lower intercostal on the left, with variable contributions from other sources, to supply most of the thoracic spinal cord and conus medullaris. Another large radicular artery from C5 or 6 usually predominates supply to the cervical cord. All radiculomedullary arteries enter the dura at the nerve root sleeve where they give off branches that supply the dura. Here, where a selective nerve root block or misguided transfemoral epidural injection is often done, is where the nidus of most spinal arteriovenous malformations is located. The space is occupied by epidural fat, which in at least one cadaveric study does not exist between the dural sac and pedicle, surrounding a valveless venous (Batson’s) plexus and varies in both thickness and configuration along the neuraxis. Although anatomically first described in the 1830s, the role of epidural venodilation as a primary cause of radicular pain has only recently been reported. Studies with mixed results based on Sicard and Forrestier’s loss-of-resistance technique have largely focused on the accuracy of placing an injectate within this space with contrast-enhanced video fluoroscopic confirmation of successful placement. Variable data exist with the accuracy of non-image-guided injection improving with easily palpable anatomic landmarks, but there is no information on vascular uptake.

Many vascular structures, both pathologic and normal variants, can exist within or border this potential space. Only recently have the arterial and venous patterns of the human dorsal root ganglion with a largely internal arterial and superficial venous drainage been deciphered and standardized. This study concluded that the ganglion had a propensity to a relative compartment syndrome when challenged with compression. The vasculoanatomy becomes more relevant when considering structural pathology in the surrounding area, namely, HNP. Veins in this region are both prevalent and easy enough to cannulate that historically, lumbar sacral epidural venography was used as a diagnostic modality in the era before magnetic resonance imaging and found to have an 88.2% to 98% accuracy when compared with intraoperative findings. One study even found it 8% more accurate than myelography. In another study (n = 304), a positive epidurogram, either complete, incomplete or unilateral block, was found in 100% of 270 surgically verified intervertebral disk protrusions, and three negative epidurograms from the same series showed no pathology on surgical exploration. However, it should be noted that clinical improvement has never been shown to correlate with improvement on epidurogram during lysis of epidural adhesions.

The importance of avoiding unintentional vascular injection cannot be overstated. Six case reports preceded a 1982 presentation to the U.S. Food and Drug Administration evaluating 33 cases of cardiac collapse from accidental intravascular bupivacaine, including 22 cases of maternal cardiac arrest with 15 fatalities. Complications such as hypotension, local anesthetic toxicity, total spinal anesthesia, and dural puncture are recognized complications, as are the risks of the adverse effects of preservatives, transient blindness, and pneumocephalus in the case of unintentional intrathecal injection. It is reasonable to consider that the “minor complication rate” with lumbar epidural injection, reported as high as 9.6%, is at least partially due to a degree of vascular uptake. Evidence that some degree of vascular uptake does occur with epidural injection is cited by venous air microbubble emboli being consistently seen in the systemic circulation by transesophageal echocardiography within 15 secs of the introduction of 5 ml of air into the epidural space after negative aspiration. This should not be surprising when one considers intraspinal air has been found in association with right main stem bronchus laceration and as a presenting sign of occult pneumothorax after blunt chest trauma. Part of the reason for these findings is that the vascular anatomy of the epidural space is not only more difficult to image but may be less consistent than previously believed. Contrast-enhanced magnetic resonance angiograms in myelopathic patients with normal digital subtraction angiography have shown medullary veins arising from the median veins and coronal venous plexus on the cord surface to communicate with the epidural venous plexus. This study alone challenges the notion that the epidural space and spinal cord proper have independent vasculoanatomy. Silicone rubber injection techniques have shown the thoracic posterior venous plexus becoming more voluminous and even sinusoidal in the lumbar region. Although not identically comparable with the CEC data in this study, both the CEC data and silicone rubber injection study are
supportive of a related study showing a greater preponderance of vascular uptake at the S1 neural foramen (21.3%) when compared with lumbar transformaminal (8.1%) epidurals. Unpredictable vascular pathology, such as cavernous malformations, traditionally thought of as angiographically occult, exists along a spectrum that may involve the vertebral body with or without epidural extension, solely the epidural space, or intradural extramedullary space in addition to being purely intramedullary. A purely epidural cavernous hemangioma has been reported to originate from a nerve root and supposedly extend into the ipsilateral lung apex, and extension of potentially vascular mass lesions from the perivertebral space proper to the epidural space is common. These malformations could not be more clinically relevant to intervention spine procedures when considering only mild rotatory movement with an up to 1.7 cm translation, solely the epidural space, or intradural extramedullary space in addition to being purely intramedullary. A purely epidural cavernous hemangioma has been reported to originate from a nerve root and supposedly extend into the ipsilateral lung apex, and extension of potentially vascular mass lesions from the perivertebral space proper to the epidural space is common. These malformations could not be more clinically relevant to intervention spine procedures when considering only mild rotatory movement with an up to 1.7 cm translation.

Even with a technically accurate spinal angiogram costing greater than $50,000 and requiring over 4 hrs per case, the existence of angiographically occult spinal dural fistulae is well documented. Vascular patterns having negative aspiration and negative passive fill may simply be the result of technical error such as needle movement, or a complicated multifactorial interplay of anatomy (vessel diameter), physiology (venous pressure), and physics (vessel collapse on aspiration). Pressures within both the epidural space and venous system fluctuate with physiologic variables, such as phase of respiratory cycle, position, and fluid status. This last situation is more relevant in the larger subpopulation of chronically ill patients, such as patients with chronic renal failure, congestive heart failure, diabetes, or vasculopathy, who are poor or nonsurgical candidates and therefore more likely to undergo spinal injections as part of conservative management. This is especially important when considering that most symptomatic paravertebral arteriovenous malformations are the result of compression by epidural veins or conges
tive myelopathy. Limitations of current imaging capability becomes even more evident when considering cases such as a surgically verified idiopathic encapsulated hematoma, which had clinically, radiographically (computed tomography and rim-enhancing lesion on magnetic resonance imaging with gadolinium), myelographically, and discographically (leakage of contrast material into the mass) been diagnosed as an L5/S1 extruded disk. Computed tomographically guided injections are performed at some institutions, and although there may be an advantage in obtaining intrarticular needle placement in arthritic facet joints, real-time vascular uptake is not readily obtained with this imaging modality.

**CONCLUSION**

Ample data based on anatomic and radiographic studies exist to suggest that there is a wider range of variability and pathology in dural, epidural, paravertebral, and spinal cord vascular anatomy than is traditionally believed. Imaging studies are imperfect, and most are static images of dynamic structures or challenged by postural and physiologic variables for which compensation cannot be predictably made. This study demonstrates that neither aspiration for blood nor delay and observation for passive filling are adequate measures to rule out intravascular injection. There may be a statistically lower rate of intravascular injection by translaminar approaches with theoretically greater consequences in the cervical spine. Spinal injection will never be risk free, but the safest method of injection is fluoroscopically guided, contrast-confined injection, which should be considered the current standard of care.

**REFERENCES**


Habitual Physical Activity Levels Are Associated with Biomechanical Walking Economy in Children with Cerebral Palsy

ABSTRACT

Objective: To evaluate in children and adolescents with cerebral palsy the relationship between habitual physical activity and biomechanical treadmill walking economy and whether treadmill belt speed or walking time affect economy.

Design: Physical activity was measured in 11 subjects (10.6–16.3 yrs) with mild cerebral palsy using a triaxial accelerometer. To determine biomechanical walking economy, subjects’ stride lengths and vertical sacral excursions were measured during each minute of three 3-min walks on a treadmill (at 60%, 75%, and 90% of individually determined fastest treadmill walking speed).

Results: Biomechanical walking economy at 60%, 75%, and 90% of (their) fastest speed each explained about half of the intersubject variance in daily physical activity (movement counts). A similar relationship was found between these biomechanical walking economy variables and movement counts at or above the 80th and 90th percentile (total minutes per day, number of 5-min bouts per day). Walking economy was 23.9% higher when subjects walked at 90% than when they walked at 60% of their fastest walking speed. No other speed-related effects on economy were found, nor did time affect economy.

Conclusions: Within this population, those with high biomechanical treadmill walking economy are the more habitually physically active. Treadmill belt speed, but not walking time, affects biomechanical walking economy.

Key Words: Cerebral Palsy, Physical Activity, Gait, Child, Adolescent
Cerebral palsy (CP) refers to a “group of nonprogressive, but often changing, motor impairment syndromes secondary to lesions or anomalies of the brain arising early in development.”1 It occurs 2–2.5 times per 1,000 live births, and spastic diplegia and hemiplegia2 are the most common subtypes. One manifestation of the motor impairment is low biomechanical walking economy (mechanical energy conservation). Mechanical energy expenditure (EE) of those with mild CP (able to walk without support) during treadmill walking at 3 km/hr was found to be 50–70% higher than controls, with the variation due to the method for calculation of mechanical power.3 Children with CP also have low habitual physical activity (PA). Compared with controls, a group of ten subjects with CP, nine of whom were ambulatory, demonstrated a 15% lower physical activity level (PAL).4 The role of biomechanical walking economy in the level of habitual PA in CP, however, is unknown.

A meaningful investigation of the relationship between biomechanical walking economy and habitual PA requires that both variables be quantified appropriately. There is, however, no agreement in the literature on an optimal, global measure of biomechanical walking economy. Unnithan et al.3 used a segmental dynamics approach to calculate total body mechanical EE during treadmill walking in the absence of kinetic data from a force plate. With this method, the kinetic and potential energy of body segments is estimated based on kinematic and anthropometric data, with or without various assumptions about energy transfer within and between body segments. Total body mechanical EE, calculated in this manner, however, can vary greatly, depending on which assumptions are made.3

In the mid-1990s, Kerrigan et al.5 introduced the biomechanical efficiency quotient (BEQ). The BEQ (which we consider a biomechanical economy quotient because efficiency is not specifically assessed) is the ratio of the measured vertical sacral excursion during a gait stride to that predicted from sacral height and stride length data. Individuals who are healthy with a mature gait pattern would therefore have a BEQ of 1. The equation for the predicted value is based on a mathematical model the authors6 developed from the kinematic gait data of able-bodied adults. The underlying assumptions are that the sacrum approximates the center of mass and that mechanical energy conservation is maximized when vertical center of mass excursion is minimized.7 The BEQ, unlike the segmental dynamics model, avoids assumptions regarding the inertial characteristics of body segments and energy transfers. In adults and children with various neurologic impairments, the BEQ is sensitive to changes in walking proficiency. The percentage change in BEQ due to the use of ankle foot orthoses correlated ($r = 0.73$) with the percentage change in comfortable walking speed.5 Pilot data from our lab have shown the between-day reliability of the BEQ during treadmill walking to be high ($R > 0.90$) in both healthy adults and in children and adolescents with mild CP.

Habitual PA in CP has been objectively quantified by estimating total metabolic EE using the doubly labeled water technique4 and the heart rate (HR) FLEX method.8 To investigate the role of biomechanical walking economy in these individuals’ habitual PA, a PA measure that is based on metabolic EE may not be appropriate, however, as metabolic EE is not simply related to kinematic events. Lower limb antagonist muscle coactivation, for example, is related to metabolic EE in this population, at least during treadmill walking.9 To investigate the role of biomechanical walking economy in habitual PA, a mechanical measure of PA, such as can be obtained from an accelerometer, may be more appropriate. Accelerometers are well accepted for use in PA studies with typically developing children10 and have also been used to assess PA in young people with physical disabilities.11

The primary purpose of this study was to determine the relationship between biomechanical walking economy, as measured by the BEQ, and the mechanical aspect of habitual PA, as measured by accelerometry. Because children and adolescents with mild CP are able to walk at different speeds both over ground and on a treadmill,9,12 we also chose to examine the effect of treadmill belt speed on the BEQ and minute-by-minute BEQ differences within a walk. We hypothesized that there would be a positive, linear relationship between habitual PA and biomechanical walking economy. We also hypothesized that the subjects’ walking proficiency would be less at more demanding, faster speeds and, thus, that biomechanical walking economy would be lower at faster compared with slower speeds. Because 12–15 mins of treadmill walking practice seems sufficient for habituation from a metabolic walking economy perspective,12 we hypothesized that there would be no minute-by-minute differences in the BEQ within a walk at any speed. It was assumed that the subjects in this study would be habituated to walking on a treadmill according to published procedures.12
METHODS

Subjects

Seven boys and four girls, 10.6–16.3 yrs, with mild spastic CP, level I or II as determined by from the Gross Motor Function Classification System, participated in the study (Table 1). Eight of the subjects had walked on a treadmill within the previous year. Five of them used the handrails; four did not. One subject regularly walked on a treadmill once per week (without holding on to the handrails). None of the subjects had orthopedic surgery within the previous year or had taken medication to reduce spasticity within the preceding 6 mos. Two subjects habitually wore hinged ankle foot orthoses during the time they participated in this study. The subjects had a similar degree of soft-tissue contracture. They had at least 105 degrees of hip flexion and lacked no more than 20 degrees of hip extension. With the hip extended, passive range of motion at the knee was full. With the hip at 90 degrees, the subjects on average lacked about 40 degrees of knee extension. Ankle dorsiflexion with the knee extended was at least 0 degrees (neutral). Subjects were otherwise healthy and receiving no medication that would affect the variables measured in the study. All subjects had at least one parent or guardian with full-time employment and lived within a 15-min drive of an urban center. All families had a vehicle. No subject was a member of a competitive sports team during their time in the study. Eight of the subjects reported that they were physically active 1–2 hrs/wk in unorganized PA (games with neighborhood friends, walks with family members). Three of them reported they were active on average for a similar amount of time per week in organized activities such as horseback riding or exercising at a gym. They refrained from caffeine for 3 hrs, eating for 2 hrs, and heavy exercise for 8 hrs before coming to the lab for each visit. Written, informed consent was obtained from subjects ≥14 yrs of age. For subjects <14 yrs old, written, informed consent was obtained from a parent, preceded by ascent from the child. The study was approved by the McMaster University Research Ethics Board. Subjects were recruited through local children’s rehabilitation centers.

Design

All subjects were tested during the early school year (late summer to midautumn). The subjects visited the lab on two occasions and were visited in their homes in the morning and evening of three different days (two weekdays and one weekend day). If they habitually wore ankle foot orthoses, then they continued to do so while in the study. During the first lab visit, they were habituated to walking on the treadmill without holding on to the handrails, and their fastest treadmill walking speed (FWS) was determined. During the second lab visit, they walked on the treadmill 3 mins at each of 60%, 75%, and 90% (0.8 m/sec) of their previously determined FWS. The lab and home visits were completed within 16.4 ± 21.1 and 12.9 ± 6.2 days, respectively. Subjects were in the study for 27.6 ± 19.6 days. Each lab and home visit lasted 2–2.5 hrs and 25 mins, respectively. The order of testing was randomized when test order could affect results.

Measurements

Lab Visit 1

The subjects completed questionnaires regarding habitual PA (modified from Bar-Or14), health status, and diet (time, content, amount of the last meal and snack), with a parent’s help if necessary. Pubertal stage (pubic hair for boys, breast development for girls) was self-determined, based on photographs15 according to the criteria of Tanner.16 Total body length, body adiposity, and limb dominance were determined as previously described.12 Body mass (Mott Electronic Scale UMC1000, accuracy ± 10 g; Ancaster Scale, Brantford, ON, Canada) was measured after subjects emptied their bladder, while subjects wore shorts, socks, and a T-shirt. To subsequently calculate total body mass and nude body mass (Table 1), all clothing, with the exception of underwear but including shoes and braces (if used) and equipment worn during

<table>
<thead>
<tr>
<th>TABLE 1 Subjects’ characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjects</td>
</tr>
<tr>
<td>Age, yrs</td>
</tr>
<tr>
<td>Body mass, kg</td>
</tr>
<tr>
<td>Sum of four skinfolds, mm</td>
</tr>
<tr>
<td>Arm span, cm</td>
</tr>
<tr>
<td>Pubertal status</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Topographic distribution of spasticity</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>GMFCS</td>
</tr>
<tr>
<td>Lower limb dominance</td>
</tr>
<tr>
<td>GMFM-E, %</td>
</tr>
<tr>
<td>CWS, m/sec</td>
</tr>
<tr>
<td>FWS, m/sec</td>
</tr>
<tr>
<td>Surgical history</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

GMFCS, Gross Motor Function Classification System; MAS, Modified Ashworth Scale; GMFM-E, Gross Motor Function Measure Dimension E (Walking, Running and Jumping); CWS, comfortable walking speed (over ground); FWS, fast walking speed (over ground); SDR, selective dorsal rhizotomy; HL, hamstring lengthening; BI, bilateral; HAL, hip adductor lengthening; UNI, unilateral; TAL, tendo-achilles lengthening; Mean (SD) listed for group data.
the treadmill walks were also weighed (Accuba Scale 1200, accuracy ± 0.10 g). The same clothing and footwear were worn for both lab visits. Classification of the topographic distribution of spasticity was based on Minear. One person (D. B. Maltais) determined severity of gross motor involvement using the Gross Motor Function Classification System, a five-level system in which level I denotes those with the mildest involvement. To screen for contractures, lower limb passive range of motion was assessed (goniometry) using standardized techniques modified from McDowell et al. Gross motor function related to walking ability was measured using the Walking, Running, and Jumping Motor Function Scale, a five-level system in which level I denotes determination of gross motor involvement.

Walking ability was also assessed by measuring the comfortable and fast walking speed on level ground (30-m walkway) using the median of a triplicate measurement. Subjects rested (sitting) between each trial until HR (Polar Vantage XL, Polar CIC, Port Washington, NY) was within 10% of its preexercise value. The subjects were taught how to walk on a recently calibrated treadmill (Woodway Desmo M Tread Erogometer, Woodway USA, Waukesha, Wisconsin) without holding on to the handrails, and their FWS on the treadmill was determined. Two subjects, who like the other subjects, walked over ground without support, stabilized themselves by supporting themselves slightly and periodically during all treadmill walks with their fingers on the handrails (alternating hands).

Lab Visit 2

The subjects walked on the treadmill three times at each of 60%, 75%, and 90% FWS. They rested in between walks in standing (supporting themselves with their arms) for 5–8 mins until breathing was not labored and they indicated that they were ready to begin the next walk. To calculate biomechanical walking economy, vertical sacral excursion and stride length during each gait cycle, during each treadmill walk, were measured using a rear-mounted, previously calibrated, kinematic data acquisition system (OptoTrak, Northern Digital, Waterloo, Canada). Sampling rate was 50 Hz. The accuracy of measurement in our lab has been previously determined to be 0.2 mm. Four nonco-linear infrared markers were mounted on each of three lightweight thermoplastic shells. One shell was affixed at the sacrum on a wide elastic belt that the subjects wore, and one shell was affixed at back of each shoe over the heel. These latter two shells were formfitting. All wires from the infrared markers converged in a lightweight junction box that was affixed to the back of a snug-fitting vest that the subjects wore. A cable from the junction box was affixed to the back of a snug-fitting vest that the subjects wore. A cable from the junction box was connected to a telemetric controller unit worn on the belt. A rechargeable battery that was also attached to the belt provided battery power to the infrared markers.

Home Visits 1–6

Subjects were visited in their homes in the morning (after the subject had risen, and dressed) and evening (before the child undressed for bed) of three different days (two weekdays and one weekend day). If the child took physical education at school or was involved in organized PA, one of the PA monitoring days was a day when these events took place (organized activities usually took place on the weekend). During the morning visit, the PA monitoring equipment was secured on the subject. Activity was monitored with a triaxial accelerometer (RT3, Triaxial Research Tracker, Stayhealthy, Eklander IA). All subjects were monitored with the same unit. The accelerometer in the RT3 is a single chip sensitive to acceleration in three orthogonal axes (X, Y, Z, representing movement in the vertical, anteroposterior, and mediolateral directions, respectively, when the monitor is worn as directed by the manufacturer). Acceleration is measured along each axis over user-defined epochs (1 sec or 1 min). The data are stored as “activity counts.” The relationship between acceleration (1 g = 9.81 m/sec) and activity counts, however, is unclear. The user can also set the device to record a vector magnitude (vector sum of the activity counts from the three axes) at each epoch. In this study the RT3 was set to measure vector magnitude activity counts and record them summed over each minute (activity counts = vector magnitude activity counts). The RT3 was worn in a felt pouch on a belt at the waist and secured with tape over the dominant hip, under the clothing. No swimming or bathing took place during the data collection periods, so there was no need to remove or reapply the equipment during the day. Start time was noted on a watch synchronized to the internal clock in the accelerometer. In the evening of each monitoring day, the subjects were again visited and the equipment was removed. The subjects were interviewed about what they did during the day. Whether and when the subjects climbed stairs or rode in a vehicle and for how long was also recorded. We also asked the subjects what they did before and after the monitoring each day, when they were not sleeping. The evening visit lasted about 25 mins. In total, PA was monitored for 11 hrs daily (from about 8 a.m. to 7 p.m. on weekdays and from about 9 a.m. to 8 p.m. on weekends). Subjects slept 8–10 hrs per night.

The RT3, a recently available triaxial accelerometer based on the technology of the previ-
viously available Tritrac R3D (Reining International, Madison, WI), was used because it provides a reliable and clinically appropriate measure of habitual PA. The older R3D has been shown to more accurately predict the oxygen uptake of children during typical activities compared with the uniaxial WAM accelerometer. When adults walk and run on a level treadmill, the newer RT3 predicts oxygen uptake similarly to the R3D. The RT3 is smaller (71 × 56 × 28 mm, 65.2 g) than the R3D (120 × 65 × 22 mm, 168 g), which may make it more appropriate than its predecessor for quantifying PA in children. Pilot data from our lab with subjects with mild CP (n = 11), whose PA was monitored simultaneously with the RT3 and HR over two weekdays and one weekend day, showed a high, positive correlation (r = 0.88) between the PAL, the ratio of total EE to resting EE, and the activity counts per day as measured by the RT3. Reliability (intraclass correlation) of the RT3 during mechanical vibration tests meant to simulate low and low-moderate PA levels is high (r = 0.99) indicating a strong relationship between axis across frequencies and across the monitors (n = 23). Intramonitor variability across frequencies and axis, also evaluated in this previous study, showed a coefficient of variation of 0.9–15%, with the coefficient of variation decreasing as frequency increased. Similar findings with the RT3 were reported during treadmill walking and running.

Calculations and Data Reduction

Kinematic Data

Using custom software written by one of the authors (M. R. Pierrynowski), gait cycles were manually determined from the kinematic data from one infrared marker on the right foot. Motion data from the four infrared markers was collected to give data redundancy in case of marker “drop out” due to obscuring of an infrared marker, such as by the contralateral limb during the swing phase of gait. Within a subject, the same infrared marker was used to determine all gait cycles. All cycles beyond 3 SD of the mean cycle length for a trial (0–1.3%) were considered atypical and were discarded. Stride length for all remaining gait cycles was calculated automatically by the software using initial contact (the lowest position of the heel during the gait cycle) as the start and end points of the stride. The sacrum, as an estimate of center of mass, was modeled by the lowest of the three infrared markers that formed a downwardly pointed triangle on the sacral shell. Again, the extra infrared marker and subsequent data redundancy at the sacrum were to ensure that the motion of the sacrum was captured throughout each gait cycle. Vertical sacral height was the distance from the lowest of the sacral markers to the treadmill belt on which the subject was standing. Standing sacral height for each stride was estimated automatically by the software as the average vertical height of the sacrum during the stride. In this manner, individual differences in the placement of the sacral shell relative to the actual sacrum were minimized. Measured vertical excursion of the sacrum during each stride was calculated by the software as the difference between the highest and lowest vertical sacral positions (relative to the treadmill belt) during the stride. Only strides for which all heel data points and all sacral data points were available were used in the calculation of the BEQ. The predicted vertical excursion of the sacrum for each gait cycle was determined from the following equation:

\[ p = \frac{1}{2}(h - \frac{1}{2}\sqrt{h^2 - (\frac{1}{4}l)^2}) \]

where \( h \) was the sacral height and \( l \) was the stride length. The BEQ was expressed as m/p, where \( m \) was the measured vertical sacral excursion and \( p \), the predicted value. A BEQ was calculated for each gait cycle. An average BEQ over minutes 1, 2, and 3 for each walk was also calculated.

PA Data

The sum of the minute-by-minute vector magnitude activity counts (11 hrs each day) was averaged over the three monitored days to yield an activity count per day. The total number of minutes (averaged across the 3 days) that the activity count was at or above the 80th and 90th percentiles for the group was also determined. From graphs of the minute-by-minute activity count data, the total number of blocks of 5, 10, and 20 mins in which the activity counts were higher (11 hrs each day) was averaged over the three monitored days to yield an activity count per day. The total number of minutes (averaged across the 3 days) that the activity count was prolonged (5, 10, 20 mins) at or above these percentiles. The 80th percentile was an activity count of 375, which for most subjects corresponded to walking at a slow or comfortable walk within a room, according to the interview information. The 90th percentile was an activity count of 736, which corresponded to faster walking, that occurred when subjects reported to be walking with friends or walking outside. For times when the subject reported to be in physical education class or taking part in other sporting activities, the activity counts were higher (>1000).

Statistical Analyses

Simple linear regression was used to assess the relationship between each of the biomechanical walking economy variables (the BEQ at 60%, 75%, and 90% FWS) and each of the habitual PA vari-
ables (activity counts per day, number of minutes per day activity counts were at or above 80th and 90th percentile, number of 5-, 10-, and 20-min bouts per day activity counts were at or above 80th and 90th percentile). The effect of treadmill belt speed (60%, 75%, and 90% FWS) and treadmill walking time (minute 1, minute 2, minute 3) was analyzed using a 2-way, repeated measures analysis of variance. Tukey's honestly significantly different post hoc test was used to identify relevant pairs that were significantly different. Alpha was set at 0.05. All analyses were performed using Statistica for Windows (Version 5.5, StatsSoft, Tulsa, OK).

RESULTS

Some kinematic data loss occurred during the treadmill walks as the BEQ calculation program automatically rejected a cycle if there was any marker drop out. For eight subjects, 98–100% of their typical strides for each treadmill walk went into the calculation of the BEQ variables. For the other three subjects, about half of their walks had minimal data loss (99–100% of their typical strides were used in the BEQ calculations). Of the remaining walks, all but one had at least 20 strides that went into the BEQ calculation. For one walk, for one subject, seven strides were available with no marker drop out, with at least two strides used to calculate the BEQ for each minute of the 3-min walk. With respect to the accelerometer data (movement counts), there was no data loss due to mechanical failure of the accelerometer. Nine of the 11 subjects wore the accelerometer for all 11 hrs of each of the three monitored days. The other two subjects wore the accelerometer for the first 8.5 hrs of one of the weekdays and for the full 11 hrs of the other weekday and on the weekend day. Because the interviews from these latter two subjects showed that their activities (within subject) were similar during each of the weekday evenings monitored, the activity count data from the same time on the other weekday for each subject was used to estimate their PA during the 2.5-hr period when they did not wear the accelerometer.

Because the subjects' BEQ at 60%, 75%, and 90% FWS were highly intercorrelated ($r > 0.98$), the relationships between the BEQ at each speed and a given PA variable were very similar (Fig. 1, Table 2). The subjects' BEQ at 60%, 75%, or 90% FWS explained 54%, 52%, and 53%, respectively, of the variance in total activity counts per day ($r = -0.73, -0.72, -0.73$, respectively, $P = 0.01$). Figure 1 shows that lower activity counts per day were associated with higher BEQ values (lower biomechanical walking economy). A similar relationship was found between the BEQ at each of the three speeds and the total number of minutes per day at or above the 80th and 90th percentile (Table 2). There was no significant relationship between the BEQ at any of the speeds and the number of 10- or 20-min bouts per day at or above the 80th and 90th percentile (Table 2). Independent of time, the BEQ was lower when the subjects walked on the treadmill at 90% FWS than when they walked at 60% FWS (mean difference, $0.97 \pm 0.7$ 95% confidence interval, $P = 0.03$, Fig. 2). There were no other interspeed differences in the BEQ. The BEQ was not affected by time at any speed, nor was there any interaction between treadmill walking speed and time (Fig. 2).

DISCUSSION

The main findings from this study with children and adolescents with mild CP were as follows: (1) biomechanical walking economy explained a large proportion (50%) of the variance in habitual

FIGURE 1 Relationship between activity counts per day and biomechanical walking economy at 60%, 75%, and 90% fastest walking speed (FWS). BEQ, biomechanical economy quotient; *$P = 0.01$. 
TABLE 2 Correlations for relationships between the biomechanical economy quotient and daily physical activity at or above the 80th and 90th percentile for the group. Group mean (standard error of the mean) for activity count data also shown.

<table>
<thead>
<tr>
<th>Activity Counts at or Above 80th Percentile</th>
<th>Activity Counts at or Above 90th Percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minutes per day</td>
<td>Minutes per day</td>
</tr>
<tr>
<td>BEQ 60%</td>
<td>BEQ 75%</td>
</tr>
<tr>
<td>No. of 5-min bouts per day</td>
<td>-0.73&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>No. of 10-min bouts per day</td>
<td>-0.55</td>
</tr>
<tr>
<td>No. of 20-min bouts per day</td>
<td>-0.43</td>
</tr>
</tbody>
</table>

BEQ, biomechanical economy quotient; 60%, 75%, 90%, percentage of fastest treadmill walking speed.

<sup>a</sup><em>P</em> < 0.05.

PA, with the exception of prolonged PA (10- and 20-min bouts), (2) biomechanical walking economy was higher when the subjects walked at 90% than when they walked at 60% FWS, and (3) there were no minute-by-minute differences in biomechanical walking economy at 60%, 75%, or 90% FWS.

Our hypothesis that there would be a positive, linear relationship between habitual PA and biomechanical walking economy was, for the most part, confirmed. Subjects with low biomechanical walking economy or high BEQ values, showed low habitual PA, which is a low value for the number of activity counts per day, number of minutes per day activity counts were at or above the 80th and 90th percentile, and number of 5-min bouts per day activity counts were at or above the 80th and 90th percentile. The lack of a significant correlation between the BEQ at any speed and more prolonged habitual PA, 10- and 20-min bouts per day with activity counts at or above 80th and 90th percentile, likely reflects the similar low levels of prolonged activity in most subjects (Table 2) rather than a failure to find the correlations significant due to low statistical power. Low levels of prolonged moderate to high-intensity PA are also typically seen in healthy children and adolescents of an age similar to those in the present study. With the present sample size of 11, with alpha set at 0.05 and beta set at 0.20 (statistical power = 0.80), there was sufficient statistical power to detect positive or negative correlations of about 0.70 or stronger (i.e., correlations where at least half the variance in habitual PA could be accounted for). Correlations (positive or negative) of <0.70 are considered moderate at best and may be considered suspect. An examination of the curves suggested by the data in Figure 1, however, suggests there may be asymptotes at each end. If this were to be the case, then for subjects with very high and very low BEQ values, there would no longer be a relationship between their BEQ and habitual PA. These potential, nonlinear relationships, however, are based on only one or two data points in the present study and thus could be unstable. Further research with larger sample sizes is needed if more complex relationships are to be explored.

It is difficult to determine from the movement count values reported in this study whether the subjects with CP had low levels of habitual PA compared with typically developing children and adolescents. The study was not designed to answer this question or test such a hypothesis. In addition, there are no data in the literature for RT3 activity counts for those with CP or for typical developing children or adolescents. As noted in the “METH-ODS,” however, PAL values (total EE/resting EE) were calculated for the present subjects using HR data simultaneously collected along with activity counts and individual HR-oxygen uptake calibrations (not reported) done in the lab. The mean HR-based PAL for the group was 1.37 ± 0.18. These values are lower than the 1.56 ± 0.18, calculated using doubly labeled water in mostly ambulatory (eight of nine walked), but younger children (mean age, 8 yrs) with spastic diplegic CP. PAL (also calculated using doubly labeled water) for typically...
developing children and adolescents by contrast, is higher (1.7–2.0).26 Between-study differences in the age of the subjects with CP and the methods for calculating PAL make direct comparisons between the studies difficult. The data, however, suggest that although the subjects in the present study have mild CP, their habitual level of PA is likely compromised. These subjects' BEQ values, on the other hand, are all >1, indicating reduced biomechanical walking economy.5 The BEQ values in this present study seem to reflect values for subjects with a neurologic impairment, but whose biomechanical walking economy is not as severely reduced as some of the subjects in the previous study. This probably reflects the milder nature of the neurologic impairment of the present subjects compared with those in the previous study. All subjects in the previous study, for example, wore ankle foot orthoses, whereas only two subjects did in the present study. Whether walking over ground compared with on a treadmill affects the BEQ is unknown.

Although we found a strong relationship between biomechanical walking economy and habitual PA, it remains unknown whether low biomechanical walking economy results in low habitual PA. Bar-Oz27 has suggested that low walking economy, might be one of the causes of these subjects' early fatigability.28 If that is the case, then these subjects may self-regulate their level of PA to avoid fatigue. Further research is needed to determine whether interventions used to improve walking proficiency in this population also result in improved biomechanical walking economy and if improved biomechanical walking economy corresponds to an improvement in habitual PA. It is also possible that higher levels of habitual PA result in improved biomechanical walking economy. Exercise programs, which presumably increase habitual PA (at least temporarily), have been shown to improve the metabolic economy of locomotion of typically developing children.29 Research determining the effect on biomechanical walking economy of increasing habitual PA, perhaps through an exercise intervention program, would help to clarify whether low biomechanical walking economy might in part be the result of low habitual PA.

We hypothesized that biomechanical walking economy would be lower at faster compared with slower speeds. This hypothesis was not confirmed. From the data, it is not clear why biomechanical walking economy was higher at the fastest compared with the slowest speed. When subjects similar to the present group walked at 90% FWS, their oxygen uptake per stride was (found to be) significantly higher than when they walked at 60% FWS.12 This indicates that subjects with mild CP are more metabolically economical when walking at the slower rather than the faster speed. When individual subjects in the present study are considered, however, there is much variability in the effect of treadmill walking speed on the BEQ. Four of the subjects showed the hypothesized pattern, their BEQ was lower at 60% compared with 90% FWS. One subject showed no difference in BEQ at these two speeds. The remaining six subjects showed the group pattern, with their BEQ higher at 60% compared with 90% FWS. The effect of speed on the BEQ seems to be associated with the subject’s biomechanical walking economy. Subjects’ mean BEQ averaged across the two speeds was highly and negatively correlated ($r = -0.97, P = 0.0001$) with the difference in BEQ between 90% and 60% FWS (BEQ at 90% FWS − BEQ at 60% FWS). A more complete answer to the question of the effect of treadmill belt speed on BEQ is that the effect is highly related to the subject’s BEQ. From the equation predicting this relationship ($y = -0.652x + 1.3571$) it seems that the transition BEQ value is just over 2 (2.08). Subjects with a mean BEQ above this value may be more economical at 90% FWS and subjects with a mean BEQ below this value may be more economical at 60% FWS.

Interestingly, post hoc statistical analysis showed that the subjects’ speeds at 90% FWS were not significantly different than their comfortable walking speed over ground, whereas their speeds at 60% FWS were significantly slower (0.26 ± 0.4 m/sec). It seems that as a group, these subjects could not walk on the treadmill much faster than their comfortable overground walking speed.

Because a 12- to 15-min treadmill walking practice session seems sufficient for habituation from a metabolic walking economy perspective,12 we hypothesized that there would be no minute-by-minute differences in the BEQ within a walk at any speed. This hypothesis was confirmed. Our finding of no minute-by-minute differences in biomechanical walking economy agree with the finding of no minute-by-minute differences in metabolic walking economy, after steady state is reached at minute 2 of a 4-min walk.3 The subjects with CP in this previous study were similar to the present subjects and were habituated to walking on
the treadmill in a manner similar to the procedures used in the present study.9

One possible limitation of this study is the large variability in the ages of the subjects because habitual PA generally decreases with age.24 When the relationship between habitual activity (activity counts per day) and age was assessed post hoc, however, these two variables were not significantly correlated ($r = 0.33, P = 0.32$). Similarly, there was no significant relationship between the other habitual PA variables and age. Thus, there is likely no bias in the habitual PA data due to age. Age was also not significantly related to the subjects’ mean BEQ ($r = 0.39, P = 0.23$).

Another study limitation is that because the BEQ is a global measure of biomechanical walking economy, the particular aspect or aspects of the subjects’ walking that contribute to their low biomechanical walking economy is unknown. In healthy adults, the major determinant than minimizes vertical sacral excursion seems to be heel rise. Heel rise from foot-flat raises the sacrum when it is at its minimum height and thereby reduces its vertical excursion by up to 66%, compared with what vertical sacral excursion would be with a theoretical compass gait.30 Further research is required to assess whether heel rise has a similar role for subjects with CP or whether other factors are major determinants of minimization of vertical sacral excursion.

In conclusion, children and adolescents with mild CP who have a high biomechanical treadmill walking economy are the more habitually physically active. Treadmill belt speed, but not walking time, also affects biomechanical walking economy. Further research is required to determine (1) activity counts from the RT3 accelerometer that reflect habitual PA in healthy children and adolescents, (2) differences in the BEQ of both typical subjects and those with mild CP during over-ground compared with treadmill walking, (3) the effect on the BEQ and habitual PA on interventions designed to improve walking proficiency or habitual PA in this population, and (4) what aspects of the walking pattern are the major determinants for minimization of vertical sacral excursion in those with CP.

ACKNOWLEDGMENTS

We thank the volunteers and their families for their participation in this study; C. Cassidy, M. Kubacki, P. Millar, R. Trott, and D. Zbogar for assistance with data collection; and M. L. Schmuck for assistance with data analysis.

REFERENCES

23. Powell SM, Rowlands AV: Intermonitor variability of the
Predicting Outcomes after Hip Fracture Repair

ABSTRACT


Objective: To compare the activities of daily living before and after hip fracture and construct a statistical model for discharge destination and independent walking. The classification accuracy of the model was determined from an independent sample.

Design: Prospective study: FIM™ prefracture, at discharge, and at 6-mo follow-up were obtained from 63 patients who underwent operations for acute hip fractures. A statistical model for discharge destination and independent walking was made and classification accuracy was checked using 78 independent samples.

Results: The motor FIM scores at prefracture decreased significantly at discharge ($P < 0.0001$) and at 6-mo follow-up ($P < 0.0001$), but at 6-mo follow-up, they had increased significantly compared with those at discharge ($P = 0.0103$). A mobility subscale was used to predict discharge destination, and mobility and social cognition subscales were related to independent walking. The predictive accuracy was 87%.

Conclusions: Motor FIM scores increase for at least 6 mos after hip fracture, and discharge destination and independent walking were highly predictable from FIM mobility and social cognition subscales.

Key Words: Hip Fracture, Activities of Daily Living, Statistical Model, Predictive Accuracy
Hip fractures are common among the elderly and greatly affect the activities of daily living (ADLs). It is useful for both the patients and family to know, as soon as possible, the likely functional recovery and discharge destination after hip fracture. Numerous reports have examined the ADLs after hip fracture and used these to predict the likely functional recovery and discharge destination of hip-fracture patients.1-13 In most cases, however, the intention was not to develop statistical models but to simply state the factors associated with a particular outcome.1 Kitamura et al.2 suggested that the age, walking ability, and ADLs before fracture and at 2 wks after surgery, dementia, and a history of contralateral hip fracture are significant predictors of patient ambulatory ability at 1-yr follow-up. Similarly, Parker and Palmer3 stated that prefracture mobility, mental state, physical health, age, and type of fracture were significant factors in predicting the ability to continue living at home at 1-yr follow-up. In addition, Cheng et al.4 suggested that age and preoperative ambulation level affect the ambulation progress, whereas Koval et al.5 noted that a patient aged ≥85 yrs was predictive of failure to recover the basic ADL status enjoyed prefracture. Ceder et al.6 stated that the ability to walk 2 wks postsurgically, living with someone, general medical condition, and type of fracture were important for a direct return home. Some authors have also suggested that cognitive status is not an important factor for motor gain,7,8 whereas others state that it is strongly associated with functional recovery.9,10

Some statistical models for predicting postsurgical outcomes have previously been presented,3,6,10,11 however, the clinical usefulness of such models is only known when they are applied to a sample independent from that which they were developed.1 To the best of our knowledge, Duke and Keating1 are the only authors to have developed statistical models and shown their accuracy with regard to hip-fracture patients using an independent sample. They stated that the maximum distance walked and ability to transfer from a supine to sitting position on postsurgical day 2 are significant predictors of transfer independence and ambulation at postsurgical week 2; they achieved an 88% classification accuracy from the independent sample. Although their data are useful in understanding ADLs at postsurgical week 2, we also need to know the ADLs after this time. Magaziner et al.9 suggested that almost maximum recovery in all functional domains is achieved by 6 mos, with little additional recovery observed at the population level between 6 and 12 mos. Therefore, a follow-up of at least 6 mos after surgery is thought to be useful.

The main surgical and rehabilitation goal for hip-fracture patients is the reacquisition of ambulation. Whether they are able to return to their own home is also important. Rehabilitation centers specifically catering in hip fractures are uncommon in Japan; therefore, patients tend to remain in the hospital longer than in Western countries, with possible discharge destinations being their own home or a nursing home. The purpose of this prospective study was (1) to compare the ADLs before and after hip fracture, (2) to make a statistical model for discharge destination and independent walking, and (3) to show the classification accuracy of the statistical model from an independent sample.

**METHODS**

The ADLs of patients who had hip fractures were evaluated by motor and cognitive FIM™ subscales,14 which include 18 items. Each item uses a 7-point ordinal scale from total assistance (a score of 1) to complete independence (a score of 7). The motor FIM scores are divided into four subscales that include 13 items as follows: self-care (eating, grooming, bathing, upper body dressing, lower body dressing, and toileting), sphincter control (bladder management and bowel management), mobility (bed, chair and wheelchair transfer, toilet transfer, and bathtub or shower transfer), and locomotion (walk or wheelchair and climb stairs). The potential scores of these items range between 13 and 91. The cognitive FIM subscale, on the other hand, has two subscales that include five items as follows: communication (comprehension and expression) and social cognition (social interaction, problem solving, and memory). Scores can range between 5 and 35. The total FIM score range is 18–126. With the walk-or-wheelchair item in the locomotion subscale, only walk was considered. When a patient was wheelchair bound daily and total assistance was needed for walking, the FIM score of this item was 1. FIM assessment was performed by trained evaluators.

**Group A**

A total of 74 consecutive patients aged ≥50 yrs were operated on for acute hip fractures in our hospital between January 1998 and December 1999. FIM scores just before hip fracture were assessed retrospectively from interviews with the patients, their family, or nursing home caregivers. FIM scores at admission were not evaluated because most of the patients could not ambulate, irrespective of their prefracture walking ability. In this setting, nursing home means a private facility for the care of individuals who do not require hospitalization but who cannot be cared for at
home. FIM score at discharge and at 6-mo follow-up were also scored by assessors who were blinded to the FIM scores at prefracture. Informed consent was obtained from the patients, their accompanying relatives, or legal guardian. Prefracture FIM scores were not available for four patients, four patients were lost to follow-up, and three died within 6 mos of surgery. Thus, 63 patients (11 men and 52 women; mean age, 78 yrs; range, 54–93 yrs; 34 right-side fractures and 29 left) who completed 6-mo follow-up were included in group A. Of these, 21 had femoral neck fractures (14 were treated by hemiarthroplasty and seven by multiple pins) and 42 had intertrochanteric fractures (28 were treated by intramedullary nails and 14 by compression hip screws).

The following FIM subscales were used for discharge destination and independent walking analysis: self-care, sphincter control, mobility, locomotion, communication, and social cognition. The patient age, the length of stay (LOS) in hospital, living status (alone or with someone, and at own or at a nursing home), type of fracture, presence of dementia, presence of heart disease, and history of stroke were also analyzed. Independent walking was defined as a score of 6 (modified independence) or 7 (complete independence) in the FIM walking item.

**Group B**

A total of 80 consecutive patients aged ≥50 yrs were operated on for acute hip fractures in our hospital between January 2000 and December 2001. The mobility and social cognition FIM subscale scores just before hip fracture were assessed retrospectively from interviews with the patients, their family, or nursing home caregivers. Informed consent was obtained from the patients, their accompanying relatives, or legal guardian. Data were not available for two patients. No patients died during their stay in the hospital; therefore, 78 patients (23 men and 55 women; mean age, 78 yrs; range, 50–94 yrs; 35 right-side fractures and 43 left) were included in group B. Of these, 29 had femoral neck fractures (23 were treated by hemiarthroplasty and six by multiple pins), and 49 had intertrochanteric fractures (47 were treated by intramedullary nails and two by compression hip screws). Seven patients were lost to follow-up within 6 mos, and 71 patients completed the 6-mo follow-up. Predictive accuracy, sensitivity, and specificity were calculated using the model generated from group A.

**Statistics**

Wilcoxon’s signed-rank test was applied to assess the differences in FIM scores at prefracture, discharge, and 6-mo follow-up. Correlations were performed using Spearman’s rank-correlation test. The unpaired t test was used to compare the LOS between groups A and B, and the χ² test for independence was used to assess the differences in discharge destination and independent walking between the groups. These analyses were performed using the statistical package StatView 5.0 for Macintosh (SAS Institute, Cary, NC). Univariate logistic regression analysis was performed to determine the influence of each discharge destination and independent walking predictor. Significant predictors were subjected to stepwise multiple logistic regression analysis, and the probability of home discharge and independent walking were obtained. The partial correlation coefficients of the selected predictor variables were also calculated. A probability of >0.5 was considered a prediction of home discharge or independent walking at discharge or at the 6-mo follow-up. Logistic regression analyses were performed using the statistical package SPSS for Macintosh 4.0 (SPSS, Chicago, IL). P < 0.05 was considered statistically significant in all analyses.

**RESULTS**

**Group A**

The average LOS of group A was 54 days. The median motor FIM score was 90 (35–91) at prefracture, 80 (14–90) at discharge, and 86 (13–91) at 6-mo follow-up, and the median cognitive FIM score was 35 (8–35) at prefracture, 35 (7–35) at discharge, and 35 (7–35) at 6-mo follow-up. The motor FIM score at prefracture decreased significantly by discharge (P < 0.0001) and at 6-mo follow-up (P < 0.0001), but at the 6-mo follow-up, it had increased significantly compared with that at discharge (P = 0.0103). The motor FIM score at prefracture was significantly correlated with that at discharge (r = 0.673, P < 0.0001) and at 6-mo follow-up (r = 0.813, P < 0.0001), and that at discharge was significantly correlated with that at the 6-mo follow-up (r = 0.962, P < 0.0001). In four patients, the FIM scores at the 6-mo follow-up increased to more than those at prefracture; three of these patients were stroke patients.

In 26 patients, prefracture FIM scores were completely independent (FIM score of 126), but none of these patients had FIM scores of 126 at discharge and 14 had scores of 124 or 125. At the 6-mo follow-up, six patients had FIM scores of 126 and 18 had scores of 124 or 125. The patients with FIM scores of 124 or 125 used canes or handrails when walking or climbing stairs; they could walk without canes but preferred to use them for safety. The other ADL items were completely independent. Therefore, patients whose prefracture FIM scores
were completely independent usually achieved good ADLs, and >90% of these patients could walk without canes at the 6-mo follow-up.

All patients who had lived in nursing homes before hip fracture (16 patients) returned there at discharge. The discharge destinations of the 47 patients who had lived in their own homes before hip fracture were predicted utilizing logistic regression analysis. In univariate analysis, self-care, sphincter control, mobility, locomotion, communication, social cognition, and dementia were significant predictors, but when entered in stepwise multiple logistic regression analysis, only the mobility FIM subscale was selected. The best predicted model was determined as follows:

**Probability of home discharge =**

\[
\frac{1}{1 + \exp[5.35 - 2.28 \cdot M]} \quad (1)
\]

where M is the sum of the three mobility items, the scores of which range from 0 to 3. The three items in the mobility subscale are equivalent to 1 if the FIM scores of bed, chair, and wheelchair transfer is ≥6, toilet transfer is ≥6, or tub or shower transfer is ≥6; otherwise, they are 0. For example, when the FIM score of bed, chair, and wheelchair transfer is 6, toilet transfer is 5, and tub or shower transfer is 4, M is 1. When the mobility of a patient before hip fracture is fully independent, M is 3, and the probability of home discharge is 82%. On the other hand, when the mobility score is fully dependent, M is 0 and the probability of home discharge is 0.5%.

In the univariate analysis, the significant predictor variables for independent walking at discharge and at 6-mo follow-up were self-care, sphincter control, mobility, locomotion, communication, social cognition, living status (own or nursing home), and dementia. In stepwise multiple logistic regression analysis (Table 1), the mobility and social cognition FIM subscales were significantly related to independent walking both at discharge and at 6-mo follow-up. The partial correlation coefficient between the mobility and social cognition subscales was −0.03.

The best predicted models were determined as follows:

**Probability of independent walking at discharge**

\[
= \frac{1}{1 + \exp[12.39 - 2.11 \cdot M - 2.54 \cdot S]} \quad (2)
\]

**Probability of independent walking at 6 months follow-up**

\[
= \frac{1}{1 + \exp[11.86 - 3.78 \cdot M - 1.19 \cdot S]} \quad (3)
\]

Each mobility subscale item is equivalent to 1 if the FIM scores of bed, chair, and wheelchair transfer is ≥5, toilet transfer is ≥5, or tub or shower transfer is ≥5; otherwise, they are 0. M ranges from 0 to 3. The social cognition FIM subscale has three items and each is equivalent to 1 if the FIM scores of social interaction is ≥6, problem solving is ≥6, or memory is ≥6; otherwise, they are 0. S is equivalent to the sum of the three social cognition items, which range from 0 to 3. When the mobility and social cognition subscales are fully independent, both M and S are 3, and the probabilities of independent walking at discharge and at 6-mo follow-up are 82 and 96%, respectively. When the mobility and social cognition subscales are fully dependent, both M and S are 0, and the probabilities of independent walking at discharge and at 6-mo follow-up are 0.0 and 0.0%, respectively. Independent walking is impossible.

**Group B**

The average LOS of group B was 41 days, which is significantly shorter than that of group A \((P = 0.0007)\). However, the discharge destination and independent walking status at discharge and at 6 mos were not significantly different between groups A and B. The 26 patients who lived in nursing homes before hip fracture all returned there at discharge. The predictive accuracy of the discharge destination of the remaining 52 patients who lived in their own homes before hip fracture was 87%, and the sensitivity and specificity were

<table>
<thead>
<tr>
<th>TABLE 1 Multivariate logistic regression analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent Variable</strong></td>
</tr>
<tr>
<td>------------------------</td>
</tr>
<tr>
<td>Home discharge</td>
</tr>
<tr>
<td>Independent walking at discharge</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Independent walking at 6-mo follow-up</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
93% and 43%, respectively (the specificity of discharge destination for group A was also low, at 50%). The predictive accuracy of independent walking at discharge for the 78 patients in group B was 87%, and the sensitivity and specificity were 87% and 87%, respectively. The predictive accuracy of independent walking at the 6-mo follow-up of the 71 patients who completed the 6-mo follow-up was 87%, and the sensitivity and specificity were 88% and 86%, respectively.

DISCUSSION

In this study, the motor FIM scores showed significant differences at prefracture, discharge, and 6-mo follow-up, whereas the cognitive FIM scores did not change. Other studies have also reported that overall FIM score improvements are primarily the result of improved motor FIM scores. Our results indicate that motor FIM scores increase for at least up to 6 mos after surgery, although they are still lower than those at prefracture. The motor FIM scores at prefracture, discharge, and 6-mo follow-up were significantly correlated, suggesting that patients with better motor FIM scores at prefracture tend to regain better motor FIM scores after hip fracture. In particular, patients with good motor FIM scores at discharge also have good motor FIM scores at 6-mo follow-up because of the high correlation between them \( r = 0.962 \). More than 90% of the patients whose prefracture ADLs were completely independent could walk without canes at 6-mo follow-up; however, many of these patients used canes because they did not want to sustain another hip fracture.

The mobility FIM subscale was the only predictor for home discharge, and the mobility and social cognition subscales were predictors of independent walking at discharge and at 6-mo follow-up. We used FIM subscales as potential predictors instead of the individual FIM items because the individual items in one subscale were usually dependent on other items in the subscale, having a high partial correlation coefficient. Although the FIM subscales are not fully independent of each other, the mobility subscale is independent from the social cognition subscale because the partial correlation coefficient of these scales was almost 0. It is possible to say, therefore, that both subscales were independent predictors for independent walking. Disagreement might exist, especially in the cognitive FIM score, but in this model, we only need to know whether the cognitive FIM score is \( \geq 6 \). If no assistance is necessary in resolving cognitive problems, the FIM score is always \( \geq 6 \); we believe that it is not too difficult to determine whether this is \( \geq 6 \) or not.

The classification accuracy of our data in an independent sample was good, except for the specificity of discharge destination. The low specificity of discharge destination in both groups A and B suggests that patients who were placed in nursing homes at discharge were often predicted to return to their own homes. Discharge destination depends not only on medical conditions and family relationships, but also on the available healthcare resources. Japan does not have sufficient rehabilitation centers and facilities for home rehabilitation; however, the Japanese government recently attempted to shorten LOS to reduce medical costs. Kitamura et al. reported that the average LOS for hip-fracture patients was 67 days from data obtained in 1992. The average LOS in this article was 53 days from 1998 to 1999 and 41 days from 2000 to 2001. The LOS in Japan is still longer than in Western countries, but a sufficient social care system has yet to be established. Therefore, some patients give up on the idea of returning to their own homes and instead enter nursing homes.

CONCLUSIONS

Motor ADLs increase for at least up to 6 mos after hip fracture. More than 90% of the patients whose prefracture ADLs were completely independent could walk and climb stairs without canes or handrails at the 6-mo follow-up. Some patients even have a chance to achieve better ADLs than existed prefracture. The mobility subscale was attributed to predict discharge destination, and the mobility and the social cognition subscales were related to independent walking both at discharge and at 6-mo follow-up. Our model is simple because only one predictor was used for discharge destination, and just two were used for independent walking; very important attributes for clinical use. Although our sample was small, one advantage of our model is that classification accuracy was obtained from an independent sample. The model for independent walking was proven to be useful because of its high percentages of predictive accuracy, sensitivity, and specificity. However, caution is needed in the use of our model for discharge destination because of its low specificity. A better model could be expected to be obtained from a multicenter study using a larger sample.

ACKNOWLEDGMENTS

We thank Masato Takahama, MD, DMSc, and Toshiki Kuroda, MD, DMSc, at the Department of Orthopaedic Surgery, Akita City Hospital, for their cooperation.

REFERENCES


Influence of Aquatic and Weight-Bearing Exercises on Quantitative Ultrasound Variables in Postmenopausal Women

ABSTRACT

Objective: In this prospective, controlled study, the effects of weight-bearing and aquatic exercises on the calcaneal ultrasonic scores of postmenopausal sedentary women were investigated.

Design: A total of 62 postmenopausal sedentary women (mean age, 54.1 ± 7 yrs) with broadband ultrasound attenuation (BUA) T-score variables less than −1 were admitted to Atatürk Balneotherapy and Rehabilitation Center and randomized into aquatic exercise (n = 21), weight-bearing exercise (n = 21), and control (n = 20) groups. The subjects were told to perform the aerobic exercises according to the Borg scale. Quantitative ultrasound variables, BUA, and speed of ultrasound were evaluated after the 6-mo training study.

Results: Calcaneal BUA increased in aquatic exercise and weight-bearing exercise groups by 3.1% and 4.2% (P < 0.05, P < 0.05) respectively. There was a decrease in BUA by 1.3% in the control group (P > 0.05). Speed of ultrasound did not change in the aquatic exercise, weight-bearing exercise, or the control groups. There were no statistically significant differences between the exercise groups for BUA and speed of ultrasound. The percentage changes in the aquatic exercise and weight-bearing exercise groups were statistically significant when compared with the control group for BUA (P < 0.01, P < 0.01) and speed of ultrasound (P < 0.05, P < 0.05).

Conclusions: Although weight-bearing physical activity is known to be superior to non-weight-bearing activity to increase the bone mass, our present evidence shows that aquatic and weight-bearing exercises both can increase calcaneal BUA.

Key Words: Aquatic, Exercise, Osteoporosis, Quantitative Ultrasound, Weight Bearing
Osteoporosis due to estrogen deficiency in postmenopausal women is a significant health problem with a prevalence of 30–50%, and the prevalence of vertebral and hip fracture is rising as a consequence. In 2010, it is estimated that there will be 32 million postmenopausal women in the United States, and by the end of this decade some 5–10 million women will be diagnosed as having osteoporosis by clinical observations. Fractures of the spine and hip are known to be the major determinants affecting quality of life in elderly people. In addition, the annual expenditure of osteoporotic fractures is substantial and will increase with the age of the population.

The role of regular exercise in the maintenance of good health is receiving more attention currently than in the past. Exercise may exert a very local effect on the highly stressed parts of the skeleton or make a systemic effect, mediated by increased growth hormone (GH) secretion, which in turn increases the production of insulin-like growth factor. The mechanical strain theory predicts that a load-engendered strain will stimulate an osteogenic response until the bone becomes strong enough to bear that load with subthreshold strain. Exercise is considered to have a unique influence on bone, and the orientation of trabeculae is always reflective of the physical stresses and strains of bone. So, not only bone volume but also the distribution of the supporting bony struts within the bone volume are influenced by exercise.

A recent meta-analysis revealed that the effects of exercise on bone mass does not support for increasing or maintaining lumbar spine or femoral neck bone mineral density (BMD) in premenopausal women who do resistance exercises. However, some longitudinal, case-controlled studies support the notion that exercise increases the bone mass. In addition, weight-loaded exercises and regimens of vigorous aerobic and strength training have been correlated with bone density. Smith et al. investigated the efficacy of 4 yrs of weight-bearing exercise in deterring bone loss. A total of 80 exercise subjects exercised three times a week, 45 mins per session, and it was found that bone loss rates in the exercise subjects were lower than in the control subjects. Although regular physical activity has an influence on bone mass, the level of regular physical activity described in some of the studies is not applicable and standardized for everyone. Chow et al. offered aerobic and strengthening exercises three times a week with free weights, Kudlacek et al. preferred a senior dancing program for 3.2 hrs/wk, and Tsukahara et al. showed that a water exercise program more than once a week had anabolic effects on bone. However, in terms of identifying programs that offer public health benefits, it is important to determine whether adherence to such programs is adequate, in addition to assessing whether interventions are effective. Resistance programs that are feasible for elderly persons in research trials may not be practical in the long term because of the unknown cardiovascular compliance of the subjects. The fact that subjects enjoy the group activities and aerobic compliance of the middle-aged or elderly people with unknown cardiovascular tolerance determine the adherence to the programs.

Research into the effects of exercise interventions on the trabecular thickness and orientation of the weight-bearing bones make quantitative ultrasound (QUS) a current issue. QUS measures the combination of elasticity, structure of trabeculae, and density, which will provide more sensitive information about fracture risk than techniques that reflect bone density alone. QUS in vivo was first proposed in the evaluation of bone mass by Langton et al. in 1984. It is an alternative method for noninvasive assessment of skeletal status. The technology is easy to perform, free of ionizing radiation, and portable, in contrast to dual-energy x-ray absorptiometry devices, which are relatively expensive, use ionizing radiation, and require patients to be referred to hospital-based activities. The calcaneus has been most frequently used as a site for measurement, with promising clinical results. In two prospective studies, it was found that the risk of hip fracture was associated with low ultrasound values and that ultrasound variables predicted hip fracture in the elderly and BMD assessed by dual-energy x-ray absorptiometry.

The study of Chappard et al. the correlation coefficients between calcaneal broadband ultrasound attenuation (BUA) and calcaneal BMD (range, 0.81–0.90) were higher than those found in previous in vivo studies. Initial clinical reports have suggested that calcaneal ultrasound does predict vertebral fracture.

Because cancellous bone turns over much more rapidly than does cortical bone, functional adaptations are likely to occur first in this more metabolically active tissue. Alternatively, the prediction of fracture risk is enhanced by measuring an additional trabecular-rich site (i.e., the calcaneus). As in our study, the os calcis is the most widely used bone for ultrasonic measurement for several reasons: (1) it has two nearly plane-parallel sides, (2) it is surrounded only by a thin layer of soft tissue, (3) it consists mainly of trabecular bone, and (4) it is also a weight-bearing bone.

In the present study, we attempted to determine which of the moderately increased physical activities, aquatic or weight-bearing exercises, is...
more beneficial for bone in sedentary postmenopausal women.

MATERIALS AND METHODS

Subjects

A prospective, controlled, longitudinal study was carried out in which the therapeutic efficacy of aquatic and weight-bearing exercises was investigated in postmenopausal sedentary women. A total of 71 postmenopausal sedentary women from the outpatients of Atatürk Balneotherapy and Rehabilitation Center were elected according to their physical activity level and calcaneal BUA (QUS) T scores as inclusion criteria. All subjects who met the study criteria were informed of the nature of the study, and a written consent was obtained. The term sedentary was functionally defined as <1.5 km of walking or <4 hrs of standing a day, and activity questionnaires were used to confirm that all subjects were sedentary before enrollment in the protocol.24 Seventy-one postmenopausal sedentary women with BUA T scores of -1 and less were randomized into three groups (aquatic exercise, weight-bearing exercise, and control groups) and enrolled in a 6-mo training study. The mean age of the subjects was 54.1 ± 7 yrs (range, 48–61 yrs). Physical examinations and routine laboratory analyses were performed to ensure that no preexisting condition would confound results. The subjects were free of preexisting cardiovascular, metabolic, and endocrine disorders. Thoracic and lumbar radiographs of the subjects were examined to define the previous compression fractures as an exclusion criteria, and it was found that none of the subjects had fractures. The groups were compared with mean age, years postmenopause, body mass index (BMI), cigarette smoking, urinary calcium, serum calcium, serum alkaline phosphatase, BUA, speed of ultrasound (SOS), and BUA T scores, and they were found to be homogeneous for all these variables at the beginning of the study (Table 1). The subjects were instructed not to take medications affecting calcium or bone homeostasis for 6 mos previous to or for the duration of the study. Average daily intake of calcium and relevant nutrients were standardized for every subject by the nutritionists. All the subjects took 1000 mg of elemental calcium a day.

The subjects were randomized into three groups called aquatic exercise (n = 24), the weight-bearing exercise (n = 24), and the control groups (n = 23). Exercisers were specifically instructed not to participate in any exercise class or organized activity outside of the study’s exercise sessions. The 23 control subjects were instructed to maintain their sedentary lifestyle for the duration.

QUS

Measurements were performed with an osteometer DTU-one ultrasound instrument (osteometer medi Tech A/S, Denmark) on the nondominant heel (usually left) after replacing the heel in the container of demineralized water with the help of a transducer in lateral projection. All scans and analyses were performed by a single technician who was blinded to subjects’ treatment. The water in the DTU-one instrument was warmed to 30°C before the measurement. A region of interest was determined with a half-automatic system (it can be directed manually by the user when necessity arises) to compare the results and to repeat the measurements at the same point. A region of interest is defined as the area in calcaneus where the attenuation is lowest and the amplitude is highest.18 In the screening of the heel with the DTU-one, a computer image was obtained for each position. The manufacturer’s phantom was checked regularly, and there were not any changes in the

| TABLE 1 | Clinical and ultrasonic characteristics of the groups |
|----------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
|          | G1 (MV ± SD)                | G2 (MV ± SD)                | G3 (MV ± SD)                | Significancea                |
| Age, yrs | 54.28 ± 6.08                | 54.88 ± 3.85                | 55.11 ± 5.32                | NS                          |
| PMP, yrs | 7.77 ± 3.46                 | 6.44 ± 3                   | 6.83 ± 3.47                 | NS                          |
| BMI, kg/m² | 28.94 ± 4.32              | 30.98 ± 3.93                | 29.44 ± 4.63                | NS                          |
| BUA T score | -2.68 ± 1.37              | -2.06 ± 1.51                | -2.10 ± 1.15                | NS                          |
| BUA, dB/MHz | 34.5 ± 8.82               | 37.3 ± 12.45                | 38.4 ± 9.12                 | NS                          |
| SOS, m/sec | 1549.5 ± 15.32            | 1548.4 ± 13.89              | 1549.7 ± 15.21              | NS                          |
| Smoking Patients | 4                        | 7                           | 5                           | NS                          |
| ALP, units/liter | 71.82 ± 102.02        | 173.41 ± 75.65              | 155.00 ± 119.84             | NS                          |
| Urinary calcium, mmol/liter | 159.88 ± 102.02       | 173.41 ± 75.65              | 155.00 ± 119.84             | NS                          |
| Serum calcium, mmol/liter | 2.39 ± 0.10            | 2.37 ± 0.09                 | 2.24 ± 0.51                 | NS                          |

G1, aquatic exercise group (n = 21); G2, weight-bearing exercise group (n = 21); G3, control group (n = 20); MV, mean value; PMP, postmenopausal period; BMI, body mass index; BUA, broadband ultrasound attenuation; SOS, speed of ultrasound; ALP, serum total alkaline phosphatase.

aSignificance was determined at P < 0.05; There were no statistically significant differences between the groups for the variables of age, PMP, BMI, BUA, BUA T scores, number of smoking patients, serum total ALP, serum and 24-hr urinary calcium.

54 Ay and Yurtkuran

Am. J. Phys. Med. Rehabil. • Vol. 84, No. 1

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.
phantom results. A phantom attempts to emulate the in vivo measurement as much as possible in terms of geometry and acoustic properties. Presently, there are no universally accepted QUS phantoms, only “manufacturer specific” phantoms that are not anthropomorphic. This procedure would guarantee that the device’s results reflect the biological or therapeutic reality and not a device malfunction.25 The results were expressed as the SOS (m/sec) and BUA (dB/MHz). BUA and SOS are considered to be the indicator of calcaneal structure (mineral content) and elasticity respectively.18

BUA T Score

To be clinically useful, BUA results for individual patients must be related to similar values obtained from a healthy reference population. In calculation of T scores, the mean and standard deviation of the young age group (20–35 yrs) are used as the reference range, regardless of the age of the patients whose BUA is being interpreted26,27: BUA T score = (measured BUA – young adult mean BUA)/young adult standard deviation.

Physical Activity Scale

The level of physical activity was determined by a graded questionnaire that varied from sedentary to heavy vocational and avocational activity levels.24 Physical activity was categorized as related to housework, job, and sports. These categories were each rated on a scale ranging from 0 to 6, and the total score used in the analysis was defined as the sum of the three components. Walking <1.5 km/day or standing <4 hrs/day constitutes the term sedentary, and it is categorized as 0–1 according to the scale.27 Aquatic exercise and walking three times a week for 40 mins/day represent moderate physical activity in daily life.13,24

Blood Chemistry

Liver and renal function tests (serum transaminases, gamma glutamyl transpeptidase, uric acid, urea, creatinine), thyroid function tests (thyroid stimulating hormone, free T3 and T4, total T3 and T4), serum calcium, phosphate, total alkaline phosphatase, glucose, and levels of urea, creatinine, calcium, and phosphate in 24-hr urine samples were measured for each subject. Routine laboratory analyzers (Abbott Alcyon 300i, Advia Centaur) were used to exclude the presence of underlying metabolic diseases. Blood and urine samples were collected between 8:00 and 10:00 a.m. after a 12-hr fast.

Exercise Program

Intensity of the exercises in both of the groups was adjusted to a submaximal level (intensity, 10–13) according to the Borg scale, and subjects were instructed to perform the exercises between these levels.28 The exercises were performed in groups with the guidance of a physiotherapist as an exercise instructor.

According to the exercise regimen,7,13 the first-week aquatic exercisers did 5 mins of warming up (walking slowly in the water and breathing), 10 mins of aerobic exercise (walking fast, jumping, and swaying in the water), 5 mins of cooling down (walking slowly in the water and breathing), and 5 mins of stretching to iliopsoas, hamstrings, quadriceps, gastrocnemius, pectoral muscles, and dorsal extensors (outside the pool). In the second week, aerobic exercise was prolonged to 15 mins. The duration of aerobic exercise was gradually prolonged to 25 mins until the fourth week. In the fourth and following weeks, the total duration of the exercise in one session was 40 mins. The sub-maximal aerobic exercises were done three times a week. The therapeutic pool in which the aquatic exercises were performed was 4.15 m in width, 8.20 m in length, and 1.25 m in depth. The water in the exercise pool was 1.20 m deep, and the water temperature was 29–30°C. The blood pressures of the subjects were checked before and after the exercises by a sphygmomanometer. The exercise was stopped two times in every session, and the radial pulses of the subjects were measured by the help of a chronometer. The exercise regimen was the same (warming up, walking, jumping, swaying, cooling down, and stretching) for the weight-bearing exercise group, and they performed the exercises on a flat platform in the rehabilitation department. The subjects of the control group were instructed to maintain their sedentary lifestyle for the duration.

Borg Scale

Ratings of perceived exertion, according to the 6–20 scale as proposed by Borg,28 is an index of subjective exercise tolerability. According to the exercise intensity scale, 6–7 is very easy, 8–9 is easy, 10–11 is moderate, 12–13 is quite hard, 14–15 is hard, 16–17 is very hard, 18–19 is very very hard, and 19–20 is maximum.

Functional Capacity

Because the subjects were middle-aged women of unknown cardiovascular fitness, an evaluation of effort tolerance needed to precede the inception of the exercise program. The systolic pressure usually increases during the exercise. A moderate rise of the diastolic blood pressure is acceptable, but this increase should not exceed 20 mm Hg. A decrease in systolic pressure, chest pain, or dyspnea indicates that the heart is stressed beyond its capacity, and the physical activity needs to be stopped if the
subject develops chest pain, dizziness, or dyspnea. The blood pressures of the subjects were checked before, during, and after the exercises by a sphygmomanometer, and the exercise was stopped if there was an unexpected effect. The heart rates during the exercise should not increase to more than 60–80% of maximum for each subject. Maximum heart rate is roughly calculated at 220 minus age. For example, if the mean age of the subjects is 54, maximum heart rate = 220 – 54 = 166; 60% of maximum heart rate = 166 × 0.60 = 99.6 beats/min; 80% of maximum heart rate = 166 × 0.80 = 132 beats/min. Therefore, heart rates must be in the range of 100 and 132 for a 54-yr-old participant.

The exercise was stopped two times in one session, and the radial pulses and blood pressures were measured during the exercises. The mean value of the radial pulses was determined to be 110 ± 10 beats/min. The subjects with the radial pulses or blood pressures beyond the appropriate target rate were asked to rest and then rejoin the group to continue.

Statistical Analysis

The groups were compared with Kruskal-Wallis test to determine whether any differences existed among the initial mean values of the groups for age, postmenopausal period, BMI (kg/m²), BUA scores (dB/MHz), SOS scores (m/sec), serum calcium (mmol/liter), serum alkaline phosphatase (units/liter), and urinary calcium (mmol/liter). There were no statistically significant differences between the groups, and the groups were determined to be homogeneous for the clinical, laboratory, and ultrasonic characteristics (Table 1). Wilcoxon’s rank-sum test was used to determine the changes between baseline and follow-up in each group (Table 2). Finally, Kruskal-Wallis and Mann-Whitney U tests were used to put forth the differences among the groups and to indicate their statistical significance, respectively (Table 3). Because of the nonnormal distribution of the raw scores and the small sample sizes, a nonparametric test (Mann-Whitney U) was selected. Pearson’s correlation analysis was used for the assessment of the relation among BMIs and BUA scores at baseline for each group. The level of significance for all tests was $P < 0.05$.

RESULTS

During the study period of 6 mos, nine subjects were excluded who started to use medication affecting calcium and bone homeostasis (two of them started to use hormone replacement therapy and three of them started to use bisphosphates) and who failed to complete the study regardless of reason (five of them quit the exercise groups). Therefore, the aquatic exercise group, the weight-bearing exercise group, and the control group continued with 21, 21, and 20 subjects, respectively.

The percentage of changes in the BUA variable for the aquatic exercise group, weight-bearing exercise group, and control group were 3.1%, 4.2%, and −1.3%, respectively (Table 3). The changes were found to be statistically significant for both of the exercise groups ($P < 0.05$, $P < 0.05$) but not for the control group ($P > 0.05$) (Table 2). The changes in SOS scores ($P > 0.05$, $P > 0.05$, $P > 0.05$) and BMIs ($P > 0.05$, $P > 0.05$, $P > 0.05$) were not statistically significant for the aquatic exercise, weight-bearing exercise, and the control groups (Table 2).

There were no statistically significant differences in the comparison of the aquatic exercise and

### Table 2: Statistical significance of the changes in aquatic exercise, weight-bearing exercise, and the control groups for ultrasonic parameters and body mass indexes (BMI)

<table>
<thead>
<tr>
<th>Groups</th>
<th>Initial Values (MV ± SD)</th>
<th>6-mo Values (MV ± SD)</th>
<th>Significance$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>BUA, dB/MHz</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1</td>
<td>34.5 ± 8.82</td>
<td>35.6 ± 11.21</td>
<td>$P &lt; 0.05$</td>
</tr>
<tr>
<td>G2</td>
<td>37.3 ± 12.45</td>
<td>38.9 ± 9.09</td>
<td>$P &lt; 0.05$</td>
</tr>
<tr>
<td>G3</td>
<td>38.4 ± 9.12</td>
<td>37.9 ± 8.96</td>
<td>NS</td>
</tr>
<tr>
<td>SOS, m/sec</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1</td>
<td>1549.5 ± 15.32</td>
<td>1555.4 ± 13.01</td>
<td>NS</td>
</tr>
<tr>
<td>G2</td>
<td>1548.4 ± 13.89</td>
<td>1554.8 ± 9.09</td>
<td>NS</td>
</tr>
<tr>
<td>G3</td>
<td>1549.7 ± 10.54</td>
<td>1549.1 ± 15.21</td>
<td>NS</td>
</tr>
<tr>
<td>BMI, kg/m²</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1</td>
<td>28.94 ± 3.42</td>
<td>28.07 ± 4.57</td>
<td>NS</td>
</tr>
<tr>
<td>G2</td>
<td>30.98 ± 3.93</td>
<td>28.91 ± 5.93</td>
<td>NS</td>
</tr>
<tr>
<td>G3</td>
<td>29.44 ± 4.63</td>
<td>29.51 ± 1.96</td>
<td>NS</td>
</tr>
</tbody>
</table>

$^a$MV, mean value; BUA, broadband ultrasound attenuation; G1, aquatic exercise group (n = 21); G2, weight-bearing exercise group (n = 21); G3, control group (n = 20); SOS, speed of ultrasound.

$^b$Significance was determined at $P < 0.05$; the changes were found to be statistically significant for BUA ($P < 0.05$, $P < 0.05$) but not for SOS ($P > 0.05$, $P > 0.05$) in the exercise groups; in the control group, the changes in BUA and SOS were not significant ($P > 0.05$, $P > 0.05$); BMIs in G1, G2, and G3 did not change significantly ($P > 0.05$, $P > 0.05$, $P > 0.05$) in the 6-mo period.
In the study of Chappard et al., the correlation specific BUA and BMD measured at calcaneus in vivo. They also compared imaging QUS of variation of 1–4% in a study of region of interests. Roux et al. found an average short-term coefficient of variation of 1–4% in a study of region of interests. Considering the relatively small age-related changes, there was a weak positive correlation among BMIs and BUA scores for the aquatic exercise group (P > 0.05, r = 0.13) and the weight-bearing exercise group (P > 0.05, r = 0.03). In the control group, there was no significant correlation (P < 0.05, r = 0.27) among the same variables.

**DISCUSSION**

In this longitudinal study, it was determined that a moderate increase of the regular physical activity, either as aquatic or weight-bearing exercise, is effective to increase calcaneal BUA by 3.1% and 4.2%, respectively, in sedentary postmenopausal women.

The influence of several factors on the precision of calcaneal BUA was investigated. Cross-sectional studies on age-related changes of QUS variables demonstrated substantial decreases during the period immediately after menopause but also in very elderly subjects. The changes were comparable with those observed using conventional radiographic, based-bone densitometry. Considering the relatively small age-related changes, immersion time, water depth, water temperature, rotation about the long axis of the foot and dorsal-plantar translation in QUS measurements, it is essential that everything is done to optimize precision. Roux et al. found an average short-term coefficient of variation of 1–4% in a study of region of interests of different sizes. They also compared imaging QUS with BMD and found an r value of 0.88 for site-specific BUA and BMD measured at calcaneus in vivo. In the study of Chappard et al., the correlation coefficients between calcaneal BUA and calcaneal BMD (range, 0.81–0.90) were higher than those found in previous in vivo studies. The relationship between BUA and physical density has also been investigated in vitro. Bouxsein et al. found that both calcaneal BMD and BUA were highly correlated with femoral failure load in vitro (r = 0.63 and r = 0.51, respectively).

Preliminary studies suggest that QUS might be appropriate as an assessment tool to measure response to antiresorptive therapies. QUS may be a perfect tool to screen very large numbers of people in the field for osteoporosis risk. On the other hand, it is believed that peripheral measurements may not reflect changes in the axial skeleton, the places in which the fractures mostly exist. However, recent prospective fracture studies have demonstrated that BUA and SOS at the calcaneus can predict osteoporotic fracture, as can dual-energy x-ray absorptiometry at the spine and hip. QUS is able to discriminate between normal subjects and subjects with low BMD or with osteoporotic fractures. Moreover, QUS can measure the bone loss associated with aging, estrogen deficiency, and immobilization, but the usefulness of QUS in monitoring the effects of specific treatment has yet to be defined by the follow-up studies.

Moreover, the response at a particular skeletal site may depend on the type of the treatment. In fact, exercise acts mainly on trabecular bone, which is characterized by a high turnover rate. The calcaneus, which is composed almost entirely of trabecular bone, seems to be a good skeletal site to monitor the effects of aerobic exercises such as walking and jumping.

Alterations in the bone remodeling at the time of the menopause often cause rapid bone turnover and accelerated bone loss. A sedentary lifestyle is another risk factor for osteoporosis that augments the effects of postmenopausal bone loss. Weight-bearing exercise is suggested to be a dynamic mechanical stress for the skeleton. It has generally been presumed that for exercise to be effective in preventing bone loss with aging, it must be weight bearing (i.e., walking, jogging, weightlifting, tennis playing, stair climbing) in nature to generate enough mechanical strain for am-

| TABLE 3 Comparison of the groups by percentage of changes in broadband ultrasound attenuation (BUA) and speed of ultrasound (SOS) |
|---------------------------------|------------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                                | G1 (MV ± SD)     | G2 (MV ± SD)    | G3 (MV ± SD)    | G1–G2           | G1–G3           | G2–G3           |
| BUA, %                         | 3.1 ± 0.59       | 4.2 ± 0.25      | −1.3 ± 0.49     | NS              | P < 0.01        | P < 0.01        |
| SOS, %                         | 0.4 ± 0.8        | 0.4 ± 0.3       | −0.2 ± 0.5      | NS              | P < 0.05        | P < 0.05        |

G1, aquatic exercise group (n = 21), G2, weight-bearing exercise group (n = 21), G3, control group (n = 20); MV, mean value.

Significance was determined at P < 0.05; there was no statistical significance in the comparison of G1 and G2 for the percentage of changes in BUA and SOS; the comparison of G1 and G3 and G2 and G3 indicated statistical significance for BUA (P < 0.01, P < 0.01) and SOS (P < 0.05, P < 0.05).
In cross-sectional studies, regular physical activity and aerobic fitness have been correlated with bone density. In general, this positive relationship has been supported by exercise intervention trials that have included weight-bearing exercises and regimens of vigorous aerobic and strength training but not by trials that have used low-intensity exercises such as walking. However, Kudlacek et al. reported that regular physical activity, such as senior dancing (waltz, folklore dancing) for 3.2 ± 0.8 hrs/wk for 12 mos, had a certain positive effect on spinal BMD in osteoporotic patients. Thus, it is suggested that dance movements have a weight-bearing effect on muscle strength at the back that is directly related to an increase of spinal BMD. Nevertheless, the type and optimal amount of the physical exercise needed to prevent bone loss still remain unclear.

There are various studies investigating the influences of exercise on bone. Lanyon found that bone responds to the proportion and to the amount of stress placed on it. Abramson and Delagi showed that weight-bearing forces and muscle contractions generate stress on bone necessary to prevent bone loss. Sinaki et al. confirmed this finding and found a positive correlation between BMD of the lumbar vertebrae and back extensor strength with a strain-gauge dynamometer in postmenopausal women. In another study by Sinaki et al., it was demonstrated that non–weight-bearing exercise consisting of back-strengthening exercises with backpacks (containing weights equivalent to 30% of the maximum isometric back muscle strength) increased muscle strength, but it was not effective in decreasing vertebral bone loss in ambulatory, healthy, postmenopausal women. However, their exercise protocol, consisting of 10 back extensions per day, may not be enough to produce adaptive changes in the bone mass. Blanchet et al. suggested that leisure physical activity could influence QUS variables independently of BMD and that quantitative ultrasound could be a suitable outcome measure in exercise studies in postmenopausal women. Brooke-Wavel et al. demonstrated that in postmenopausal women, brisk walking three times a week for at least 20 mins was associated with an increase in calcaneal BUA, but unfortunately, no changes were seen in BMD of the calcaneus, hip, or spine, as measured with dual-energy x-ray absorptiometry. From this point of view, it cannot be taken for granted that there was any effect on the hip and spine. On the other hand, these findings suggest the possibility that loading through walking may produce structural reorganization of trabecular bone even in the absence of changes in BMD. In the study by Brooke-Wavel et al. and in the present study, it was also determined that BUA scores improved but SOS did not, which may point that bone mineral content, reflected by BUA, may be more sensitive to short periods of exercise than the elasticity property reflected by SOS.

Krolner et al. showed that 1 hr of walking twice a week for 8 mos increased bone mineral content of the lumbar vertebrae by 3–5%, whereas in the controls, it decreased by 2–7%. Therefore, the data suggest that physical exercise can inhibit or reverse involutional bone loss of the lumbar vertebrae in normal women and that physical exercise may prevent spinal osteoporosis.

On the other hand, there are only a few studies about the influences of non–weight-bearing exercise (i.e., swimming, aquatic, or water exercise and stationary cycling) on BMD. Bloomfield et al. showed that lumbar spinal densities of seven postmenopausal women exercising regularly at moderate intensities for 8 mos on bicycle ergometers revealed a significant increase. Tsukahara et al. reported that the BMD of postmenopausal women increased slightly while participating in a water exercise program. The water exercise program included light calisthenics, swaying-jumping, and walking more than once a week for 35 mos. The level of physical activity had two maximum working heart rate peaks (approximately 120 beats/min) during the 45 mins. In the exercise group, it was found that the rate of change in the BMD showed a slight increase rather than a decrease, irrespective of the duration of menopause. Malliopoulou et al. showed that aquatic exercise designed to improve flexibility, posture, and muscle strength increased vertebral bone mass and decreased the risk of new fractures. They explained this positive effect by the strains applied through the bones by muscle contractions. Orwoll et al. documented a greater bone mass in the distal radius and in the lumbar spine of male masters swimmers who had been training for 13 ± 11 yrs, compared with sedentary controls. Interestingly this finding did not carry over to the female swimmers included in the study. In a preliminary study of Ay and Yurtkuran, aquatic exercise was determined to make an anabolic effect (osteogenic response) on the bone of the women shown by the hormonal (insulin-like growth factor-1, growth hormone) and ultrasonic variables. Thus, it was put forth that aquatic exercise for postmenopausal sedentary subjects should be considered for inclusion in future exercise intervention protocols. In the study of Kelly et al., it was shown that the osteogenic response to exercise was followed by an increase in growth hormone and insulin-like growth factor-1, the potent cell mitogens. Also, in the
study of Kudlacek et al., it was reported that the osteogenic response to exercise was followed by an increase in the bone-specific alkaline phosphatase, a marker for osteoblastic activity. In view of the gain in bone mass, this shows additional evidence of increased bone formation. Because calcaneus is a weight-bearing bone and the Achilles tendon is attached to it, both gravity and muscle contractions are thought to have an effect on it while walking and jumping. Owing to its high surface-to-volume ratio, trabecular bone in the calcaneus has presumed a turnover rate of about eight times that of a compact bone, and it is highly responsive to mechanical stimuli for bone remodeling. On the other hand, experimental studies have demonstrated increases in bone density and histomorphometric measures of bone mass in female rats subjected to vigorous swimming training. A total of 28 female Sabra rats (12 wks old) were randomly assigned to exercise and control groups. Exercised animals were trained to swim in a water bath of 35°C, loaded with lead weights (2% of body weight). It was reported that bone hydration properties, bone density, bone mineral content, and serum alkaline phosphatase were higher by 36%, 3%, 10%, and 67% in the exercise group, respectively.

There are data providing evidence of a prospective nature that aquatic exercise may be effective in reversing bone loss in healthy postmenopausal subjects. It can be explained with hydrodynamic principles. Viscosity of water acts as friction or resistance when walking and jumping in the water. An isokinetic exercise model was created in which the load and the velocity was constant for the whole range of motion. It allows the muscles and the skeleton to strengthen while encouraging greater physical activity and increasing self-confidence in this section of population.

Cross-sectional evidence from other investigators remains inconclusive as to whether muscular contraction independent of weight-bearing impact forces is capable of producing an increase in bone mass. Risser et al. reported that BMD of eumenorrheic female athletes was evaluated, and it was found that mean calcaneal densities of the volleyball and basketball players were greater than those of the swimmers and nonathletes. It was concluded that the higher bone densities for athletes in vertical weight-bearing activities were consistent but the swimmers’ low bone density in the lumbar spine; the less than published values for amenorrheic runners was unexpected. However, these results may be due to the difference in quantitative bone evaluation techniques used in different studies. Furthermore, recent studies also hypothesize that mechanical forces might influence trabecular microarchitecture and other BMD-independent factors such as muscular strength and coordination, which could reduce the risk of falling and fracture. There are not any existing prospective studies in which the aquatic and the weight-bearing exercises are compared.

In this study, aquatic exercise was performed at the same magnitude of land-based aerobic activity for 40 mins/day and represents a moderate physical activity in daily life. It has been reported that a moderate level of physical activity can be graded as follows: being on feet 50–70% of the time or performing a regular set of exercises, such as jogging, walking, biking, and aerobics for ≥30 mins/day and ≥2 times/wk. Even though aquatic exercise can be considered a kind of nonloading exercise, it increased calcaneal BUA of the healthy postmenopausal subjects within this study. Our findings showed that exercise causing repeated mechanical loading in the long bones, outside of the normal physiologic range, produces significant increases in bone quality and quantity assessment variables.

There is a consistent positive correlation between BMI and BMD. The association between BMI and BMD remains strong even within normal ranges of body weight and is consistent across the different sites of BMD measured. The potential mechanisms to explain this are the increased mechanical loading of bone and the factors including the endocrine systems and growth factors such as insulin-like growth factor-1. In the study of Tsukahara et al., BMIs of the groups were between 23 and 24, less than the study groups of the present study. It was found that the BMIs and BMDs of the lumbar spine were not closely related in the exercise group. In our study groups, the BMIs were >29, and the subjects were considered as overweight. Because there was no statistically significant difference between the groups for BMIs at the beginning, it was not considered as a potential factor to change the results. Moreover, BMIs and BUA scores were not correlated in our exercise groups. In the control group, there was a weak correlation among the same variables (P < 0.05, r = 0.27) in accordance with the previous data.

The major finding of the present study is the significant ultrasonic increase of the calcaneal bone in response to the aquatic exercise regimen, which is traditionally considered nearly non–weight bearing in nature. Some postmenopausal women engage in fewer daily physical activities owing to their decreased physical capability and fitness or because of pain affecting the back and other joints. Thus, aquatic exercise such as walking, swaying, and jumping in the water provides a valuable alternative to weight-bearing activities, preparing an environment in which exercise can be done in a more pain-free and safe manner for people with established osteoporosis.
or balance and gait deficits. In addition, ultrasonic search results indicate that an aquatic exercise regimen is nearly as therapeutically effective as a weight-bearing exercise regimen on the control of calcaneal bone loss.

Several limitations to this study design should be noted. These include the small number of the subjects involved in the study and the short period of exercise. Because clinically relevant changes of bone density for prevention of osteoporosis in sedentary postmenopausal women may be achieved only after longer periods of time (e.g., 1–2 yrs), 6 mos is a relatively short study period to interpret the changes in ultrasound.41

It is concluded that aquatic and weight-bearing exercises are both determined to be effective in increasing the QUS scores of the calcaneal bone. In respect to the present study, the critical point is to increase the quantity of physical activity level to an unusual state for osteogenic stimulus, independent of the kind of exercise performed. Aquatic exercise is a suitable means by which the aged and nonswimmers can move their arms and legs easily and freely, without burdening the joints, while maintaining buoyancy with moderate physical exertion. It would therefore be sensible to offer to postmenopausal women to increase their physical activity systematically. Long-term effectiveness of aquatic and weight-bearing exercises on bone mass and fracture risk should be evaluated by means of follow-up studies, especially with synchronous measurements of calcaneal QUS and BMD of the hip and spine.

REFERENCES
8. Smith EL, Gilligan C, McAdam M: Detering bone loss by exercise intervention in premenopausal and postmeno- 

pauline women. Calcif Tissue Int 1989;44:312–21
22. Dempster DW, Lindsay R: Pathogenesis of osteoporosis. Lancet 1993;341:797–804


Functional Neuroimaging and Cognitive Rehabilitation for People with Traumatic Brain Injury

ABSTRACT


Cognitive deficits are a common consequence of traumatic brain injury. Although such deficits are amenable to rehabilitation, methods for individualizing cognitive interventions are still unrefined. Functional neuroimaging methods such as positron emission tomography and functional magnetic resonance imaging are emerging as possible technologies for measuring and monitoring the cerebral consequences of plasticity associated with brain injury and for evaluating the effectiveness of rehabilitation interventions. Functional neuroimaging may even enable more customized and efficient selection, design, or adaptation of individual cognitive rehabilitation programs. We review the current literature on functional neuroimaging after traumatic brain injury, relating these findings to cognitive rehabilitation. Overall, functional neuroimaging after traumatic brain injury has shown reliable differences in brain activity within several regions of frontal cortex, partly but not uniformly consistent with neuropsychological and structural findings in traumatic brain injury. We also outline a number of promising research opportunities for applying functional neuroimaging in traumatic brain injury settings, along with associated challenges.

Key Words: Functional Magnetic Resonance Imaging, Cognitive Rehabilitation, Traumatic Brain Injury, Functional Neuroimaging, Review, Hemodynamic
The goal of this article is to review the present state of knowledge regarding functional neuroimaging (FNI) after TBI and how this relates to TBI cognitive rehabilitation. The article proceeds in four parts. We begin by briefly reviewing common cognitive impairments associated with TBI and current rehabilitation approaches. We then describe the use of FNI to evaluate cerebral function and plasticity and how FNI might be of use in a rehabilitation setting. Next, we review in detail the applications of FNI in persons with TBI and relate those to TBI cognitive rehabilitation. Finally, we summarize points of convergence from previous sections of this review and discuss areas of opportunity associated with FNI application in persons with TBI.

**Cognitive Impairment After TBI**

Cognition is frequently disrupted after TBI, and the resulting impairments persist over time, contributing to psychosocial, educational, and vocational problems. Numerous studies of neuropsychological functioning have been reported in the literature, most commonly in individuals sustaining moderate to severe TBI. Although the effects of TBI can be variable, the most consistent neuropsychological findings are in the domains of attention, memory, and executive functioning. Attention-related deficits include distractibility, impairment of sustained attention, and dual-task performance (divided attention) problems. Although many aspects of attention may be affected, sustained attention (particularly in relatively low-demand circumstances) is commonly affected. Memory-related deficits include anterograde memory, retrograde memory, and working memory difficulties. Abnormalities in “strategic” aspects of memory are also observed, affecting the use of internal and external strategies to encode, store, and retrieve information.

Use of memory strategies is closely tied to executive functioning, another function that is commonly impaired after a TBI. The term executive function describes a diverse set of high-level cognitive processes that regulate more elementary perceptual, motor, and memory operations. These functions enable individuals to appreciate context in the environment and to use this understanding to make plans, implement strategic action, and monitor and flexibly shift behavior when it is no longer appropriate. There is increasing evidence that executive function problems after TBI are most pronounced in situations that lack external structure.

Together, these three functional domains encompass the majority of the residual deficits after TBI. They also span a large proportion of cognition and, hence, warrant the use of all possible aids in effective diagnosis, monitoring, and remediation.

**Effectiveness of Cognitive Rehabilitation**

Cognitive rehabilitation attempts to address a wide spectrum of individual cognitive processes and relationships among these processes and their bearing on everyday functioning. Two wide-ranging reviews on the efficacy of cognitive rehabilitation have been published on this topic, one by the Agency for Health Care Policy and Research ([AHCPR] now Agency for Healthcare Research and Quality), and one by Cicerone et al. The AHCPR Evidence Report, “Rehabilitation for Traumatic Brain Injury,” reviewed 600 articles that addressed efficacy of cognitive rehabilitation after TBI and then narrowed these to 114 that passed their initial screen (all controlled studies, plus uncontrolled series with information about the short- or long-term outcomes associated with cognitive rehabilitation for TBI). They concluded that among these studies, two provided reliable evidence that the use of compensatory devices (notebooks and electronic aids) improve memory, whereas another indicated that compensatory rehabilitation improved self-esteem and social relationships. Two additional studies supported the use of restorative approaches to improve immediate recall. The required level of evidence in this review, however, was particularly high and generated considerable controversy.

Cicerone et al., using somewhat less controversial standards for evidence, found the literature substantially more supportive. They identified 655 articles that addressed issues related to cognitive rehabilitation, which they narrowed to 171 studies assessing interventions for people with TBI or stroke. They found 29 class I (prospective, randomized controlled) studies, 20 of which provided “clear evidence for the effectiveness of cognitive rehabilitation for subjects with acquired TBI or stroke.” Among 64 controlled class I and II studies, all but two showed significant improvement in the group that received cognitive remediation. Since this review, a number of additional studies meeting similar criteria have been found to provide additional evidence of the effectiveness of cognitive rehabilitation in the domains of attention, memory, and executive function.

The AHCPR report posed three questions that “underlie uncertainty about the effectiveness of rehabilitation services”: (1) how should fundamental concepts such as recovery, functional status, and disability be defined, (2) how should the type and severity of the injury itself be measured, and (3) which therapies are effective, and what is the best way to match patients with treatment approaches likely to be effective for them? The third question, which is the driving force...
for this article, hints at one of the great difficulties arising when evaluating the efficacy of interventions applied to groups of people with TBI: the population of people with TBI is heterogeneous with respect to severity and localization of cerebral damage and the respective consequences of the injury.4,5,22,53 and with respect to psychological, cultural, and socioeconomic background.54 Given the particularly broad cross-section of the population that is affected by TBI, one cannot expect that patients should all be treated with a similar approach. Indeed, studies demonstrate that the outcome of specific cognitive treatments can vary considerably among subgroups (e.g., mild, moderate, and severe injury subgroups),55 among individuals,5 and as a function of treatment approach.5,6,21

For example, in the area of memory, Cicerone et al.5 found four class I studies13–16 that evaluated patients with mild or moderate memory impairment. Taken together, these four provided enough evidence to support a “practice standard” for the use of compensatory techniques in those with mild memory impairment but not in those with moderate impairment. Additional studies indicate that compensatory strategies can also be an effective approach to those with severe memory impairment, both when using specific internal strategies (e.g., visual imagery)20 or external ones (e.g., memory notebook).21 In terms of individual variability, two studies in particular suggest that details regarding specific approaches (e.g., amount of structure) for training subjects in the use of a notebook may influence individual outcome.22,23 Thus, given the evidence for differences in responsiveness to rehabilitation interventions, the challenge remains how to design and match treatment options to individual patients in the least costly, most effective, and most efficient manner possible.

**Application of FNI**

**FNI to Evaluate Cerebral Function**

To generate flexible therapeutic approaches that consider individual differences, one needs flexible approaches to cognitive testing that can take into account the neuropsychological and neurophysiologic status and makeup of the individual. Cognitive testing can provide the treatment team with important diagnostic information to help guide treatment. Neuroimaging can provide information complementary to cognitive testing. In particular, structural methods—predominantly magnetic resonance imaging (MRI) and computed tomography—can help identify and localize gross structural anomalies associated with injury, including fiber damage, hemorrhage, contusion, and necrotic tissue.56–61 These techniques have been immensely helpful in characterizing structural alterations after a head injury, but they have consistently been found to under-represent the true nature of the neurologic deficit after brain injury. Particularly in cases of mild head injury, a number of studies have found no structural anomalies in patients who remain symptomatic.

FNI methods—including positron emission tomography (PET), single-photon emission computed tomography, and functional MRI (fMRI)—offer the ability to observe dynamic changes in the neurophysiologic status of the individual during cognitive assessment and treatment and are of two general types: single-shot FNI and multi-shot FNI. Single-shot FNI provides a snapshot of the brain’s baseline functional status via measures of blood flow, perfusion, or oxygen/glucose metabolism.62,63 Such variables have been demonstrated to exhibit anomalies after TBI, and these anomalies can extend well beyond the borders of observable anatomic lesions.54 Multi-shot FNI techniques, in contrast, enable the measurement of cerebral function during states of “cognitive activation.”32 In such studies, participants are instructed to engage in one or more specific cognitive activities during scanning (e.g., memorizing word lists, counting backward by sevens, etc.). As individuals engage in different tasks, blood flow in the brain is locally increased in those regions that are most active, and hence, FNI allows the researcher to identify regions of the brain that are more or less active during specific mental states or activities. Evaluation of specific brain systems or cognitive—behavioral domains (e.g., memory or executive function) is thereby achievable, which in turn enables application of FNI to specifically targeted impairments during rehabilitation. This ability to specifically target and monitor the neural substrates of specific cognitive—behavioral domains provides a complementary view on the functional status of an individual from simple neuropsychological or structural imaging measures.

Four primary technologies have been employed to investigate the neurophysiology of TBI: electroencephalography (including event-related potentials), magnetoencephalography, PET, and fMRI. The first two techniques measure the electric and magnetic fields (respectively) associated with neuronal activity and have most commonly been used to detect functional anomalies in mild TBI for which structural anomalies are not readily apparent.65 Here, we focus on the blood flow/oxygenation measurement techniques of PET and fMRI. In PET scanning, a radioactive “tracer” is delivered to the patient, either by injection or by inhalation. This tracer is thereby introduced into the blood stream and the flow or pooling (perfusion) of the blood can be tracked by radiation-sensitive cam-
erased. fMRI is based on the fact that deoxygenated hemoglobin has different magnetic properties than oxygenated hemoglobin and that deoxygenated hemoglobin is flushed out of active brain regions.

After acquisition, raw PET and fMRI signals are preprocessed and then must be analyzed for statistically reliable differences in brain activity between task and control conditions. Preprocessing typically involves (1) image coregistration (e.g., compensation for head motion between or during scans), (2) spatiotemporal normalization (e.g., smoothing the image volumes or time series), and (3) spatial transformation to a standardized space (e.g., Talairach or Montreal Neurological Institute space\textsuperscript{66,67}). Image analysis requires a statistical test designed to establish the reliability of regionally observed signal differences. Such tests can be applied to a variety of measures, including mean or peak amplitude changes in brain activity (absolute or relative change in signal level, resulting in spatial activity maps as shown in Fig. 1 for both imaging methods) or spatial extent of activity changes between experimental conditions (e.g., experimental task vs. control task). Amplitude or spatial differences are analyzed within or between subjects or groups, or even over time, as dictated by the research question at issue.

PET measurement of fluorodeoxyglucose provides a direct measure of glucose metabolism and, thereby, neuronal energy consumption. This method, in fact, provides the best signal-to-noise measurement currently available for brain function. fMRI and PET of regional cerebral blood flow, in contrast, both detect neuronal functioning indirectly by detecting hemodynamic sequelae to neuronal activity. Importantly, methodologic differences between PET and fMRI imbue the two techniques with distinct advantages. PET scanning (1) enables quantitative measures of blood flow and perfusion, (2) provides more flexibility in subject positioning, and (3) provides the capacity to study subjects in whom MRI is contraindicated. In contrast, fMRI (4) does not involve exposure to ionizing radiation as does PET, enabling extended or detailed longitudinal studies, (5) has higher spatial resolution (the ability to detect changes in small regions), and (6) has greater temporal resolution (the ability to detect changes in brain activity over small periods of time).

Finally, it is worthwhile highlighting an important and somewhat subtle distinction between cognitive-activation–based and resting-state studies of brain function. FNI during a cognitive task has the advantage of probing the specific systems required for that task in a sensitive manner. The experimental conditions, then, are normally designed to provide a targeted test of differential engagement of the regions that comprise these systems. However, the limitation of this approach is that such cognitive activation methods provide very limited information regarding other regions of the brain. Conversely, resting-state studies with single-shot FNI, or structural imaging, enable a survey of the entire brain for functional or structural findings, whether or not they are engaged by any particular task. Thus, resting-state single-shot and structural imaging complements dynamic FNI methods nicely: the former provides a less-sensitive brain-wide survey of effects, and the latter enables sensitive evaluation of specific regions or cognitive–behavioral domains of interest.

**FNI, Cerebral Plasticity, and Rehabilitation**

FNI can provide unique insight into the question of cerebral plasticity, which “is fundamental to the concept of restorative cognitive therapy.”\textsuperscript{64} For example, FNI has the potential to help distinguish between restorative effects (wherein recovery of specific neuropsychological or regional brain functions are observed) vs. compensatory effects (wherein alternative strategies are used that rely on other abilities and brain regions to bypass impair-
Before intervention, dyslexic children exhibited little or no activation of the (normally active) posterior portion of the superior temporal gyrus during a pseudoword reading task. However, this area exhibited normal activation levels after 80 hrs of an intervention that improved reading skills, strongly suggestive of a restorative effect. Similarly, Small et al. recently studied 12 patients with acute hemiparetic stroke longitudinally over the first 5 mos of stroke recovery. Patients with good recovery showed changes in activation in the cerebellar hemisphere opposite the site of cortical stroke (i.e., the cerebellar region most directly connected to the affected primary motor cortical region), suggesting at least a partially restorative-type response. In contrast, a number of other FNI studies of hemiparetic stroke have also shown prominent recruitment in motor cortical areas ipsilateral to the paresis (that is, contralateral to the lesion), suggesting a compensatory response via activation of uncrossed motor pathways to perform tasks that typically engage predominantly crossed pathways.

By providing such insight into the effects achieved by a given therapy, FNI can help guide current and future therapy to either enhance or complement observed cerebral—behavioral effects.

It remains important to demonstrate that (1) FNI can help characterize the neurophysiologic underlying plasticity, rehabilitation, or recovery, and (2) that knowledge of such underlying neurophysiology might somehow prove useful to guide rehabilitation programs. Initial evidence of this possibility is alluded to by the stroke recovery literature. Although early FNI studies indicated that recovery from hemiparetic stroke was associated with motor cortical activity ipsilateral to a hemiparetic limb (above), more recent findings suggest that good outcome is more closely associated with re-activation of structures in contralateral pathways. The often-seen ipsilateral activity early poststroke, therefore, may be a compensatory adaptation but may also be maladaptive, or may become so over time.

FNI of neuroplasticity in this case has clear implications for stroke rehabilitation. In particular, it may turn out to be beneficial to activate contralateral brain tissue in any way possible (e.g., passive motion of affected limbs), potentially tracked by FNI, to prevent learned disuse of the contralateral pathways and reduce overactivation of ipsilateral ones (e.g., as in constraint-induced therapy). The Small et al. study described above found that cerebellar activation was an even better predictor than was the classic approach to stroke prognostication—evaluation of motor function on the impaired side. Thus, at least in the stroke population, FNI would seem to provide useful indicators of recovery when measuring the resulting change in activation of specific brain regions. Physiologic measurements of rehabilitation effectiveness—particularly measurements of the cerebral substrate of specific interventions that can be made on an individualized basis—may provide novel information for rehabilitative planning.

**FNI in Cognitive Rehabilitation**

Motivation

Before prospective investigations can take place employing FNI to help guide treatment, studies must first demonstrate associations between (1) regional cerebral activation seen with FNI, (2) treatments that change this activation, and (3) targeted outcomes. The existing literature on FNI and rehabilitation overwhelmingly addresses motor rehabilitation—particularly in stroke populations—not cognitive rehabilitation. Nevertheless, the parallels between cognitive rehabilitation—highly relevant in TBI—and motor rehabilitation are clear, suggesting that similar approaches may be effective for both cognitive and motor domains. Thus, as a parallel to the course taken in the stroke literature with motor rehabilitation, we propose in Table 1 a sequence of events to help guide and gauge the maturity of FNI-based rehabilitation evaluation. The first step, identification of critical brain regions, is well underway in the stroke literature. The identification of the cerebral effects of constraint-induced movement therapy impinges on the second step, identifying and designing interventions for rehabilitation. For stroke, the third step—monitoring or enhancing the effectiveness of a therapy based on FNI—has yet to be addressed.

Understanding pathophysiology by itself can also lead to a better understanding of the processes underlying both functional and dysfunctional activity of a patient, which is likely to influence treatment approaches. Some FNI studies have already yielded unexpected findings or otherwise lent insight into motor, cognitive, or emotional processes that might eventually influence treatment. For example, Gurd et al. demonstrated that the superior posterior parietal cortex is involved in purely semantic task switching, in addition to its

---

**TABLE 1** Three proposed stages for applying functional neuroimaging to traumatic brain injury rehabilitation

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Identify brain regions critical to recovery of function</td>
</tr>
<tr>
<td>2</td>
<td>Identify or design interventions that activate such regions</td>
</tr>
<tr>
<td>3</td>
<td>Monitor or enhance the effectiveness of the interventions</td>
</tr>
</tbody>
</table>
known role in visual and spatial task switching. Also, as already discussed, the stroke FNI literature suggests that compensatory responses contralateral to the injury—often thought to be adaptive—may instead indicate weak adaptive change or even maladaptive cerebral alteration. Such findings underscore the need for vigilance in experimentation and interpretation on the part of the TBI researcher.

**Neuroimaging of Cognitive Activation in TBI**

Although there have been relatively few cognitive activation FNI studies after TBI, they have quite consistently supported the notion of prefrontal cortical dysfunction. We review here the findings of 15 FNI studies of individuals with TBI, important attributes of which are provided in Table 2.35,77–90 In the text, we group these studies according to the task performed during the FNI scanning—vigilance and continuous performance tasks, Wisconsin Card Sorting Test (WCST), memory retrieval tasks, and working memory tasks.

In one early study, Humayun et al.80 monitored cerebral glucose metabolism while patients performed a visual (luminance evaluation) vigilance task. They found increased regional glucose metabolism in both anterior temporal and anterior frontal cortices, along with decreased regional glucose metabolism in posterior temporal cortex, posterior frontal cortex, and the left caudate nucleus, relative to matched controls. The reported cortical regions were consistent with prototypical mechanical models of head injury, wherein temporal and frontal cortices are most vulnerable to damage (especially with anteroposterior acceleration/deceleration). The frontal alterations were also associated with the patients’ reported difficulties with attention. This provided some early evidence for functional disorders in the frontal lobe of individuals with TBI.

Ruff et al.88,90 performed two regional glucose metabolism PET studies of minor brain injury (an injury resulting in a <20-min loss of consciousness) using the Continuous Performance Task—a task demanding considerable attentional resources. In the first study, five of six subjects with TBI (varying injury severity) had decreased uptake in one or both frontal lobes compared with healthy controls. In the second study, Ruff et al.88 scanned nine individuals with minor TBI and residual neuropsychological deficits while they performed the Continuous Performance Task. Considerable interindividual variability was found, but the frontal and anterior temporal cortices exhibited the most consistent alterations, namely, hypometabolism of glucose. These results again fit well with standard physical models of brain injury, wherein the prefrontal and anterior temporal cortices are most prone to damage due to the skull’s interior structure.

One experimental task of particular interest in TBI is the WCST, which is considered to be quite sensitive to frontal dysfunction.91,92 One might hypothesize, therefore, that the WCST would also be a useful neurophysiologic probe for brain function evaluation via FNI. In an initial study employing the WCST, Kirkby et al.79 used PET to examine a single subject with severe TBI and compared his results with his monozygotic healthy twin and a matched control group. Reduced regional cerebral blood flow was observed in left inferior frontal cortex, along with increased flow in the left hippocampus during the WCST. The hippocampal increase was argued to be compensatory (memory) in response to the failure to activate the prefrontal cortex. Importantly, the study provided additional evidence of particular involvement of frontal cortex in dysfunction after TBI.

In an investigation of a larger TBI cohort, Lombardi et al.81 had subjects perform the WCST during fluorodeoxyglucose PET scanning and examined the correlation between WCST perseverative errors and brain activity. They found perseverative errors to be inversely related to regional glucose metabolism in the right (but not left) dorsolateral prefrontal cortex and caudate nucleus. This provided initial evidence that FNI can provide quantifiable measures relevant to functional status of the TBI brain, but it stops short of indicating whether such FNI measures provide any supplementary diagnostic or evaluative information relevant to function. The lateralization finding is provocative and could be related to strategic or other compensatory responses to injury. However, other possible explanations exist, including the particulars of the injuries sustained by their eight patients.

Three studies have investigated memory-retrieval tasks in PET studies on subjects with TBI.77,78,89 Levine et al.78 compared one subject with severe TBI and isolated retrograde amnesia against four subjects with moderate or severe TBI with little or no retrograde amnesia. The subject with retrograde amnesia demonstrated decreased activation in the right middle frontal gyrus and increased activation in a number of nonfrontal areas on the left during a cued recall task. Ricker et al.77 examined blood flow during recall of verbal word lists in five subjects with severe TBI and four healthy control subjects. Participants in this study learned lists of 12 words from three categories (fruit, clothing, tools) during five nonscanned learning trials. They were then scanned during free recall, category-cued recall, and word recognition. Results indicated significant changes in blood flow in the left prefrontal cortex in all three conditions.
<table>
<thead>
<tr>
<th>First Author</th>
<th>Year</th>
<th>Technique</th>
<th>Experimental Design</th>
<th>Hypothesis</th>
<th>Task</th>
<th>Inj. Location</th>
<th>Injury Severity</th>
<th>Control Subjects</th>
<th>Injury Severity</th>
<th>Months Since Injury</th>
<th>Age(s) at Study, yrs</th>
<th>Age(s) at Injury, yrs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Humayun, MS</td>
<td>1989</td>
<td>FDG PET</td>
<td>Cross-sectional</td>
<td>Significant differences in regional cerebral glucose metabolism of patients relative to controls.</td>
<td>Vigilance task</td>
<td>TBI (N/S)</td>
<td>3</td>
<td>healthy matched controls</td>
<td>Mild-moderate</td>
<td>3–12</td>
<td>27, 28, 40</td>
<td></td>
</tr>
<tr>
<td>2 Ruff, RM</td>
<td>1989</td>
<td>FDG PET</td>
<td>Cross-sectional</td>
<td>Examination of alterations in cerebral function in the absence of observable CT/MRI neuropathology</td>
<td>Continuous performance task</td>
<td>TBI (variable)</td>
<td>6</td>
<td>healthy controls</td>
<td>Mild-severe</td>
<td>29</td>
<td>37 ± 12</td>
<td></td>
</tr>
<tr>
<td>3 Ruff, RM</td>
<td>1994</td>
<td>FDG PET</td>
<td>Cross-sectional</td>
<td>Functional brain imaging can detect neurological deficits after minor TBI where structural imaging cannot.</td>
<td>Continuous performance task</td>
<td>TBI (variable)</td>
<td>9</td>
<td>healthy controls</td>
<td>Mild (LOC &lt;20 mins)</td>
<td>9</td>
<td>46 ± 15</td>
<td></td>
</tr>
<tr>
<td>4 Kirkby, BS</td>
<td>1996</td>
<td>PET</td>
<td>Case study</td>
<td>Cerebral reorganization after brain injury even with respect to monozygotic twin control.</td>
<td>Wisconsin card sorting test; sensorimotor control</td>
<td>TBI patient with a monozygotic twin (anterior frontal horns)</td>
<td>TBI (variable)</td>
<td>1</td>
<td>healthy controls</td>
<td>Severe (LOC &gt;17 days)</td>
<td>72</td>
<td>28</td>
</tr>
<tr>
<td>5 Levine, B</td>
<td>1998</td>
<td>PET</td>
<td>Case study</td>
<td>Right frontal hypoactivation during episodic memory retrieval.</td>
<td>Memory (retrieval) task</td>
<td>TBI (right ventral frontal cortex)</td>
<td>TBI (variable)</td>
<td>1</td>
<td>healthy controls</td>
<td>Severe + retrograde amnesia</td>
<td>38</td>
<td>36</td>
</tr>
<tr>
<td>6 Lombardi, WJ</td>
<td>1999</td>
<td>FDG PET</td>
<td>Cross-sectional</td>
<td>Correlations between WCST preservation and prefrontal cortex activation.</td>
<td>Wisconsin card sorting test</td>
<td>TBI (variable)</td>
<td>TBI (variable)</td>
<td>8</td>
<td>healthy controls</td>
<td>None</td>
<td>N/S</td>
<td>5–168</td>
</tr>
<tr>
<td>7 McAllister, TW</td>
<td>1999</td>
<td>fMRI</td>
<td>Cross-sectional</td>
<td>Assess cerebral activity under working memory loads in mild TBI.</td>
<td>Auditory N-back working memory</td>
<td>TBI (variable)</td>
<td>TBI (variable)</td>
<td>12</td>
<td>healthy matched controls</td>
<td>Mild</td>
<td>&lt;1</td>
<td>29 ± 10</td>
</tr>
<tr>
<td>8 McAllister, TW</td>
<td>2001</td>
<td>fMRI</td>
<td>Cross-sectional</td>
<td>Assess cerebral activity under working memory loads in mild TBI.</td>
<td>Auditory 3-back working memory task</td>
<td>TBI, including 12 from previous study (N/S)</td>
<td>TBI (variable)</td>
<td>18</td>
<td>healthy matched controls</td>
<td>Mild</td>
<td>&lt;1</td>
<td>32 ± 12</td>
</tr>
<tr>
<td>9 Ricker, JH</td>
<td>2001</td>
<td>PET</td>
<td>Cross-sectional</td>
<td>Less frontal activation during recall tasks (but not recognition tasks) in TBI patients.</td>
<td>Verbal recall tasks</td>
<td>TBI (variable)</td>
<td>TBI (variable)</td>
<td>5</td>
<td>healthy matched controls</td>
<td>Severe (LOC &gt;24 hrs)</td>
<td>38 ± 16</td>
<td>25 ± 3</td>
</tr>
<tr>
<td>10 Christodoulou, C</td>
<td>2001</td>
<td>fMRI</td>
<td>Cross-sectional</td>
<td>Examine patterns of activation during working memory task in moderate/severe TBI patients.</td>
<td>Auditory serial addition working memory task</td>
<td>TBI (variable)</td>
<td>TBI (variable)</td>
<td>9</td>
<td>healthy matched controls</td>
<td>Moderate or severe</td>
<td>51 ± 41</td>
<td>32 ± 11</td>
</tr>
<tr>
<td>11 Levine, B</td>
<td>2002</td>
<td>PET</td>
<td>Cross-sectional</td>
<td>Examine the functional neuroanatomy supporting memory retrieval after TBI in moderate/severe TBI patients.</td>
<td>Verbal encoding and retrieval tasks</td>
<td>TBI (variable)</td>
<td>TBI (variable)</td>
<td>6</td>
<td>healthy matched controls</td>
<td>Moderate or severe</td>
<td>50 ± 4</td>
<td>29 ± 6</td>
</tr>
<tr>
<td>12 Scheibel, RS</td>
<td>2003</td>
<td>fMRI</td>
<td>Case study</td>
<td>Extent of frontal tissue recruited by cognitive tasks is increased after TBI.</td>
<td>N-back working memory task</td>
<td>TBI (variable)</td>
<td>TBI (variable)</td>
<td>1</td>
<td>healthy controls</td>
<td>Severe (20–44 yrs)</td>
<td>12</td>
<td>46</td>
</tr>
<tr>
<td>13 Chen, SH</td>
<td>2003</td>
<td>PET</td>
<td>Cross-sectional</td>
<td>Examine the functional neuroanatomy during working memory task in mild TBI patients.</td>
<td>Spatial working memory</td>
<td>TBI (variable)</td>
<td>TBI (variable)</td>
<td>5</td>
<td>healthy matched controls</td>
<td>Mild</td>
<td>3–35</td>
<td>34 ± 12</td>
</tr>
<tr>
<td>14 Chen, JK</td>
<td>2004</td>
<td>fMRI</td>
<td>Cross-sectional</td>
<td>Functional abnormalities in cerebral activation exist even after mild TBI.</td>
<td>One visual and one verbal working memory task</td>
<td>TBI (variable)</td>
<td>TBI (variable)</td>
<td>16</td>
<td>healthy matched controls</td>
<td>Mild</td>
<td>1–14</td>
<td>28 ± 5</td>
</tr>
<tr>
<td>15 Kim, YH</td>
<td>2004</td>
<td>fMRI</td>
<td>Imaging pre/post-treatment</td>
<td>Cerebral reorganization occurs after rehabilitation for both stroke and motor-impaired TBI patients.</td>
<td>Fingers to thumb, making a fist</td>
<td>TBI (variable)</td>
<td>TBI (variable)</td>
<td>1</td>
<td>healthy matched controls</td>
<td>Within-patient comparisons</td>
<td>N/S</td>
<td>9–38</td>
</tr>
</tbody>
</table>

FDG, fluoro-deoxyglucose; PET, positron emission tomography; fMRI, functional magnetic resonance imaging; CT, computed tomography; MRI, magnetic resonance imaging; WCST, Wisconsin card sorting test; N/S, not significant; LOC, loss of consciousness.
in the healthy control group. By comparison, the TBI group failed to activate the prefrontal cortex during the free-recall condition and (consistent with the previous study) exhibited less extensive activation in some left prefrontal areas during the cued-recall condition. Both groups showed significant bilateral prefrontal activation during the recognition condition. Finally, Levine et al. employed verbal encoding and retrieval tasks in a group of six survivors of moderate to severe TBI who had since made a good recovery. Four PET scans were performed in each condition (encoding and retrieval), and compared with a matched healthy control group. Overall, TBI patients recruited neuronal networks similar to controls. However, TBI patients showed enhanced activity in frontal, anterior cingulate, and occipital regions, and a reduction in the hemispheric asymmetry found in controls. This last study conflicts with the previous two by finding increased prefrontal involvement rather than decreased prefrontal involvement during task performance, the reason for which is unclear, although it could be related to differing task performance levels relative to controls. The three studies nevertheless suggest altered prefrontal functioning in recall-related memory tasks, with relative normalization of brain activation in subjects with TBI as they receive additional structure at recall. These results are thus consistent with neuropsychological findings reviewed earlier, indicating that cognitive problems after TBI are more pronounced in poorly structured situations. Some people with TBI may be able to show normal levels of brain activity when they receive additional structure, which would likely be an important indicator for rehabilitation evaluation and subsequent treatment.

All of the remaining FNI in TBI investigations have employed some form of working memory task. McAllister et al. performed two related studies of individuals with mild TBI, examining brain function using fMRI. In the first study, 12 subjects with mild TBI and 11 healthy control subjects were scanned during auditory N-back working memory tasks of low difficulty (N = 0, 1) and moderate difficulty (N = 2). In this type of task, the subject's response depends on whether the current stimulus matches the stimulus that appeared N items earlier in the stimulus sequence (or matches a target stimulus for N = 0). Relative to control subjects, the TBI group showed disproportionately increased brain activation in response to increasing working memory load, especially in the right dorsolateral prefrontal and parietal cortices, even in the absence of performance level differences between patient and control groups. The absence of performance-level differences between TBI and control groups strengthens the interpretation of cerebral activity differences as related to TBI instead of merely performance characteristics. In a follow up study, the authors added a more difficult working memory condition (N = 3-back), and examined brain activity with fMRI in 18 subjects with mild TBI and 12 healthy control subjects. Again, they found that the group with TBI showed an increase in activation in right prefrontal cortex during the moderately difficult condition but failed to do so significantly during the most difficult condition.

In a recent fMRI study, Scheibel et al. published a case study of a severe TBI patient performing both an N-back working memory task and a go/no-go inhibition task. Brain activity from their patient was compared with a small control cohort of healthy control subjects. The patient exhibited impaired performance on the two tasks, along with expanded, bilateral prefrontal brain activation on both tasks relative to controls. As with previous studies, this was interpreted as compensatory recruitment of broader prefrontal regions consequent to prefrontal brain damage. The different performance levels, however, complicate this interpretation.

Christodoulou et al. used fMRI to examine brain activity during an auditory-serial-addition working memory task in nine subjects with moderate to severe TBI and seven healthy control subjects. The control subjects in this study showed the expected activations in the left prefrontal cortex during the working memory task. By comparison, the TBI group showed more regionally dispersed activation, including activity in the right hemisphere. McAllister et al. and Christodoulou et al. conclude that TBI alters the ability to activate or modulate neural activity, especially in the prefrontal cortex, during working memory challenges.

Most recently, two studies have employed working memory tasks exclusively in subjects with mild TBI. S. H. Chen et al. investigated a cohort of five subjects with mild TBI and five matched controls using PET when performing a spatial working memory task in which subjects were briefly shown a dot-pattern (200 msecs), maintained this in memory over a 3-sec interval, and then had to report whether a dot appeared at a specified location. The symptomatic subjects as a group exhibited reduced activity in dorsolateral prefrontal cortex and, when analyzed individually, exhibited expanded and additional prefrontal activations as compared with the control group. Interestingly, one asymptomatic patient and one patient in whom symptoms resolved before a second scanning session both exhibited activation patterns that were similar to the control group. Although pro-

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.
vocative, the generalizability of this last finding remains to be investigated.

Finally, J. K. Chen et al.86 scanned 16 symptomatic concussed athletes with fMRI, with most subjects having sustained multiple concussions. Reported symptoms included working memory or attention or executive function deficits. Two tasks were implemented: a visual and a verbal working memory task, wherein subjects were shown four of five stimuli with which they were familiar and subjects were asked to monitor if a test item was one of the four presented stimuli, or one from a different set of five (familiar) items. In both the verbal and visual (picture) tasks, the athletes generated weaker dorsolateral prefrontal cortical modulations to the task than control subjects. Although the direction of change differed from the studies of McAllister et al.,35,82 the finding of altered brain responsiveness in this region after TBI was supported by this study.

Overall, the existing FNI studies of TBI patients are still limited. They include a relatively wide range of tasks, cover the full spectrum of injury severity from mild to severe, and are overwhelmingly case studies or small-group (n < 10) studies. These observations, coupled with the heterogeneous nature of brain injuries in general, makes firm conclusions difficult. However, there are two main, consistent findings across these studies. First, alterations in cerebral activity in the frontal lobe were found to be essentially universal—across task, injury severity, and group size. Although this could be related to the “frontal” nature of the tasks (vigilance, WCST, verbal, and memory), the alterations in frontal activation are consistent with neuroimaging findings of subjects at rest82 and with the typical phenomenology and phenotype of TBI. The occasional mismatches in sign of the alteration (decreased activity in the Chen et al.86 study and increased activity in the McAllister et al.35,82 studies or the Levine et al.78,89 and Ricker et al.77 studies) remains to be fully investigated but may be related to task or performance differences among groups. Alterations in nonfrontal brain regions were generally found only when the implemented tasks were known to activate the particular nonfrontal structures observed (e.g., TBI-induced alterations in inferior temporal activity for a visual recognition task). Most often, the alterations involved reduced modulation in cerebral activity (from baseline) in the TBI subjects relative to controls.

Of the group-based studies, some also provided cerebral activation maps for individual subjects86,88,90 and generally found cerebral activity to be more diffuse or widespread after TBI, similar to post-stroke cerebral activation patterns.73 Altered (or multiple) strategies for performing a task might account for such diffuse responses, as could a more physical explanation related to damage to the vascular bed. Diffuse responses were found in mild TBI, however, suggesting that vascular bed damage may not be necessary to generate more widespread functional activation after TBI. It is worth pointing out that one might seek to explain the decreased modulation findings in group analyses as a consequence of just such dispersion—the robustness and amplitude of cerebral activity modulations will necessarily be lower in a group presenting with inconsistent spatial activity patterns (i.e., subjects with TBI) than in a group with more consistent activations (i.e., controls).85

**Neuroimaging Findings in TBI Rehabilitation**

All the studies in the previous section examine the dynamic, functional neurophysiologic changes after TBI. Although these findings are clearly related to TBI rehabilitation, the range of postinjury periods implies significant interindividual variability in recovery or rehabilitation state of the subjects. Thus, these studies do not address neurophysiologic changes directly associated with rehabilitation. To date, we are aware of only a single FNI study in a subject with TBI directly examining a rehabilitation intervention.87 In particular, Kim et al.87 performed a study of constraint-induced movement therapy on a subject with TBI (with postinjury encephalomalacia). The patient was scanned both before and after a 2-wk, 7-hrs/day constraint-induced movement therapy intervention. While in the scanner, the subject performed both a gross motor task (fist-making) and a fine motor task (sequential finger-to-thumb opposition) with the affected hand. Hand motor function was improved from preintervention to postintervention, and the improvement was retained 2 mos later (Fugl-Meyer assessment). Activation analysis indicated more brain activity ipsilateral to the injury before the intervention and a switch to the more typical, contralateral activation after the intervention. No investigations have yet been reported with more subjects with TBI nor with a cognitive-based therapy.

**Directions and Opportunities**

**Convergent Findings**

Although FNI findings directly related to TBI rehabilitation are limited—and especially so for cognitive rehabilitation—there are significant points of convergence evident in this review. In particular, these studies almost universally found TBI to be associated with the disruption of func-
tional brain activity in prefrontal cortical regions. Somewhat less consistently, TBI was associated with subcortical dysfunction, or dysfunction in other cortical regions (predominantly parietal and temporal cortex).35,83,85 Interestingly, this contrasts with a review of the static single-photon emission computed tomography literature, which found >55% of cases exhibiting basal ganglia disruption, compared with 23% frontal cortical disruption (the next most common site of dysfunction).62 The reversal in prominence of cortical vs. subcortical alterations in the two types of FNI is provocative and may highlight an important difference between single- and multi-shot FNI sensitivities. Essentially no studies reported alterations in brain function in the posterior cortex, despite the regular use of visual stimuli and the potential for posterior cortex damage from TBI. These observations provide indirect information about functionally intact and spared cerebral regions for those involved in TBI rehabilitation.

Research Opportunities

As discussed, direct investigation of FNI in TBI rehabilitation remains a largely unexplored domain. However, the existence of >20 studies applying FNI to stroke rehabilitation questions not only suggests significant areas of opportunity with people with TBI but also provides substantial methodologic and empirical guidance for implementing such studies. For example, the stroke rehabilitation studies employing FNI, discussed briefly above, generally fall into one of four broad categories: (1) within-subject cross-sectional studies (typically comparing a paretic limb performance with non-paretic limb performance), (2) cross-sectional studies employing a matched control group, (3) longitudinal studies imaging patients pre- and postintervention or at multiple time points during an intervention, and (4) longitudinal crossover studies. Lessons learned from these studies are numerous, but two bear mentioning with respect to TBI rehabilitation.

First, and just as with structural imaging, functional imaging at a single point in time after an injury, though useful in itself, can be greatly augmented by longitudinal studies. Longitudinal studies allow one to investigate the pattern of neural responses over time, which can (as suggested previously) provide insight into mechanisms underlying rehabilitation and providing more detail on the particularly important topic of individual differences.

Second, studies can be substantially enhanced by directly comparing brain imaging findings with an independent measure of functional recovery. This enables the researcher to identify cerebral activation patterns with recovery processes more robustly than simply comparing preintervention and postintervention scans.

Although the detailed findings from the stroke studies cited above remain to be fully explored and understood, they are nevertheless valuable lessons for rehabilitation in general. For example, in motor rehabilitation, it has been suggested that passive movement therapy—which activates the crossed sensorimotor pathway—could be used to speed or enhance recovery by preventing disuse-related loss of function by forcing the crossed pathway to remain active. Constraint-induced therapy, in which the patient is forced to use the affected limb,75 may indeed be an effective mechanism for achieving this goal. The extent to which the critical cerebral pathways are activated by one or the other therapeutic methodology could be monitored by FNI, enabling more timely evaluation of the effectiveness of rehabilitation.

We would argue that using the brain to operate a limb is fundamentally the same as using the brain to perform mental tasks (e.g., to remember information). Thus, cognitive rehabilitation, analogous to motor rehabilitation, requires practice “exercising” the same or different constellations of brain regions. Appropriately designed tasks, therefore, may have the same effect for cognitive rehabilitation in TBI as constraint-induced movement therapy or passive sensory activation has for motor rehabilitation. Tasks may be able to be designed to actively or passively activate (e.g.,) left or right dorsolateral prefrontal cortex, to aid in recovery from working memory deficits, just as active or passive movement is geared toward aiding in recovery of arm dexterity. This is where the three key steps in the future FNI-based research of TBI rehabilitation, which we list in Table 1, can again come into play.

Although the domain of FNI in TBI rehabilitation is still wide open for investigation, a few questions stand out as being of particular importance for individuals with TBI. We list a few of the relevant scientific questions in Table 3. Initial hypotheses to the first of these questions already exist in the form of baseline flow/volume studies, plus the FNI studies reviewed herein. Questions 4 and 5 suggest perhaps the biggest potential payoff of FNI, where treatment decisions could be based on individual variability in brain function. There exist a number of examples of using functional imaging profiles to predict responses to treatment (including medications, behavioral treatments, or surgical therapies), particularly in the neuropsychiatric literature.94–98 To our knowledge, however, no studies have yet been published that assess the ability of FNI to predict which individuals with TBI may or
may not benefit from particular types of rehabilitation programs. This is a principle focus of the authors’ present research endeavors.

Two final research opportunities, although a bit further afield, merit brief mention. First, we propose that one could combine steps 2 and 3 from Table 1, thereby potentially allowing FNI to become part of the treatment, providing biofeedback that lets the patient and clinicians know whether he or she is on the right track during cognitive (or other) rehabilitation. In this case, successful regional cerebral activation would be analogous to successful muscular activation in electromyographic biofeedback. Although this may not be practicable with PET or fMRI, less obtrusive and portable FNI technologies that have already been developed are beginning to be applied in TBI. Second, as we proposed in Table 1, it is also conceivable that FNI could be used to help design treatment strategies. If it could be demonstrated that, for instance, activation of certain regions in prefrontal areas during cognitive tasks is associated with better outcomes, then one could develop hypotheses as to which interventions were likely to be beneficial based on FNI. Some of the ethical and experimental issues surrounding hypothesis testing could then be mitigated. New or unconventional experimental approaches could be justified, in part, on the basis of FNI results that indicate a likelihood of efficacy. Research in clinical settings could develop treatments using tasks that activate the specific area associated with better outcome in people with TBI and then compare outcomes with traditional treatment. Research that demonstrates changes not only in the activities and participation dimensions but also in brain function would enhance the credibility of those outcomes and thereby help ensure efficient utilization of resources.

Challenges

Significant challenges remain for applying FNI to populations of people with TBI undergoing rehabilitation. The methodologic challenges associated with FNI measurements are substantial, but these have been discussed in detail elsewhere. In terms of tasks, stroke studies have benefited from a nearly universally applied probe task (finger-to-thumb opposition) that can be easily implemented at multiple time points for comparative evaluation. For TBI rehabilitation, no such standardized or accepted task has yet materialized (Table 2). Developing such a task, or suite of tasks, is complicated by the complexity and relatively unknown functional architecture of the prefrontal cortex, which is often the primary target of TBI. Nevertheless, a perfect standard task is not necessary, as is evident even in the stroke literature itself. The finger-opposition task is difficult or impossible for many stroke patients to perform, and it tends to elicit mirror movements, thereby complicating the interpretation of brain lateralization findings. The task has nevertheless helped provide significant information about neurophysiologic changes associated with recovery and rehabilitation.

TBI rehabilitation studies also face significant complications associated with individual variability. Such variability will be manifest in injury severity, injury location and functional consequences, comorbid variables, preinjury differences in brain structure and functional organization, and postinjury differences in treatment and rehabilitation programs. The goal is to control as many such variables as possible. In practice, strict control is seldom as feasible as desired. One particularly important future goal, therefore, should be to scan considerably larger cohorts than have thus far been investigated. Such an approach enables better generalization to the TBI population as a whole while simultaneously enabling one to parcel out differences specific to subgroups or individuals. In parallel, using relatively standardized tasks could also facilitate meta-analyses of groups in smaller studies, in cases in which such smaller studies are all that is feasible to conduct.

<table>
<thead>
<tr>
<th>TABLE 3 Scientific questions to address using functional neuroimaging (FNI) in traumatic brain injury (TBI) rehabilitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>What are the common brain regions functionally affected by TBI? Do they differ systematically based on the type of injury or location of impact?</td>
</tr>
<tr>
<td>What is the temporal profile of brain utilization in these regions over the course of recovery? Are there patterns of change specifically associated with good and poor recovery (as appears likely in the case of hemiparetic stroke recovery)?</td>
</tr>
<tr>
<td>Is cognitive recovery associated with neurophysiologic restoration of lost function, neurophysiologic compensation for it, or both?</td>
</tr>
<tr>
<td>Can FNI be used to reliably and effectively monitor progress throughout a rehabilitation program? Can it be used to help decide when to change rehabilitation therapies on an individualized basis?</td>
</tr>
<tr>
<td>Can FNI provide predictive power—over and above current predictors such as age, injury severity, and education level—in determining rehabilitation outcome after TBI?</td>
</tr>
</tbody>
</table>

Strangman et al.  
Am. J. Phys. Med. Rehabil. • Vol. 84, No. 1  

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.
Last but not least, current FNI methods (PET, MRI) are expensive, nonportable, and highly intrusive. This is in contrast to rehabilitation settings that require considerable interpersonal interaction, patient freedom, and patient motion. As mentioned above, however, emerging neuroimaging technologies are enabling portable, noninvasive, motion-robust measurements (in particular, near-infrared approaches). Such technologies open up even wider investigational possibilities, including brain imaging during rehabilitation.

Despite the challenges, the opportunities for applying FNI in TBI settings are substantial, and the potential payoff is equally substantial. Research in stroke and TBI rehabilitation can clearly be synergistic, with findings in one domain used to help interpret and spur further investigation in the other. In addition, with the development of new neuroimaging technologies, the opportunities for investigating cerebral correlates of TBI rehabilitation and for facilitating enhanced rehabilitative outcomes can only expand.

REFERENCES


35. McAllister TW, Sparling MB, Flashman LA, et al: Differen-
tial working memory load effects after mild traumatic
36. Gershberg FB, Shimamura AP: Impaired use of organiza-
tional strategies in free recall following frontal lobe damage.
*Neuropsychologia* 1995;33:1305–33
37. Shimamura AP, Janowsky JS, Squire LR: What is the role of
frontal lobe damage in memory disorders? In Levin HS, Eisen-
berg HM, Benton AL (eds): *Frontal Lobe Function and Dysfunc-
38. Levine B, Dawson D, Boutet I, et al: Assessment of strate-
gic self-regulation in traumatic brain injury: Its relation-
ship to injury severity and psychosocial outcome. *Neuro-
psychology* 2000;14:491–500
severe traumatic brain injury. *J Int Neuropsychol Soc*
2001;7:795–804
40. Lezak MD: *Neuropsychological Assessment*, ed 3. New
York, Oxford University Press, 1995
41. Goldberg E, Podell K: Adaptive decision making, ecologi-
cal validity, and the frontal lobes. *J Clin Exp Neuropsychol*
2000;22:56–68
42. Levine B, Freedman M, Dawson D, et al: Ventral frontal
contribution to self-regulation: Convergence of episodic
tive rehabilitation on outcomes for persons with traumatic
brain injury: A systematic review. *J Head Trauma Rehabil*
1999;14:277–307
44. Prigatano GP: Commentary: Beyond statistics and re-
search design. *J Head Trauma Rehabil* 1999;14:308–11
45. Cicerone KD: Commentary: The validity of cognitive reha-
ilitation. *J Head Trauma Rehabil* 1999;14:316–21
*J Head Trauma Rehabil* 1999;14:312–5
47. Cicerone KD: Remediation of “working attention” in mild
omics for the rehabilitation of memory: A randomised group
frequency on face-name recall by adults with traumatic
50. Levine B, Robertson IH, Clare L, et al: Rehabilitation of
executive functioning: An experimental-clinical validation
of goal management training. *J Int Neuropsychol Soc*
2000;6:299–312
51. Lawson MJ, Rice DN: Effects of training in use of executive
strategies on a verbal memory problem resulting from
842–54
52. Consensus conference: Rehabilitation of persons with
traumatic brain injury. NIH Consensus Development
Panel on Rehabilitation of Persons With Traumatic Brain
53. Rehabilitation of Persons with Traumatic Brain Injury.
*NHI Consensus Statement* 1998;16:1–41
54. O’Shanick GJ, O’Shanick AM: Personality and intellectual
changes, in Silver JM, Yudofsky SC, Hales RE (eds): *Neu-
ropsychiatry of Traumatic Brain Injury*. Washington, DC,
55. Mackay LE, Bernstein BA, Chapman PE, et al: Early inter-
635–41
56. Cihangiroglu M, Ramsey RG, Dohrmann GJ: Brain injury:
Analysis of imaging modalities. *Neurol Res* 2002;24:7–18
57. Bigler ED: Quantitative magnetic resonance imaging in
traumatic brain injury. *J Head Trauma Rehabil* 2001;16:
117–34
58. Newton AB, Alavi A: Neuroimaging in patients with head
59. Garnett MR, Cadoux-Hudson TA, Styles P: How useful is
magnetic resonance imaging in predicting severity and outcome in traumatic brain injury? *Curr Opin Neurol*
2001;14:753–7
60. McAllister TW, Sparling MB, Flashman LA, et al: Neuro-
61. Bigler ED: Neuroimaging in pediatric traumatic head
injury: Diagnostic considerations and relationships to neurobehavioral outcome. *J Head Trauma Rehabil* 1999;
14:406–23
brain perfusion abnormalities in mild or moderate traum-
63. Davalos DB, Bennett TL: A review of the use of single-
photon emission computerized tomography as a diagnos-
tic tool in mild traumatic brain injury. *Appl Neuropsychol*
2002;9:92–105
64. Langfitt TW, Obrist WD, Alavi A, et al: Computerized
tomography, magnetic resonance imaging, and positron
tomogram emission tomography in the study of brain trauma: Pre-
65. Gaetz M, Bernstein DM: The current status of electro-
physiologic procedures for the assessment of mild traum-
atic brain injury. *J Head Trauma Rehabil* 2001;16:386–
405
66. Talairach J, Tournoux P: Co-planar Stereotaxis Atlas of
the Human Brain. New York, Thieme, 1988
intersubject registration of MR volumetric data in stan-
dardized Talairach space. *J Comput Assist Tomogr* 1994;
18:192–205
68. Simos PG, Fletcher JM, Bergman E, et al: Dyslexia-specific
brain activation profile becomes normal following success-
activation ipsilateral to the paretic hand correlates with
functional recovery after stroke. *Brain* 2002;125(7):
1544–57
70. Chollet F, DiPiero V, Wise RJ, et al: The functional anat-
63–71
of functional MRI to assess cerebral activation of motor
function after poststroke hemiparesis. *Stroke* 1998;29:
112–22
study of subjects recovered from hemiparetic stroke.
*Stroke* 1997;28:2518–27
73. Calautti C, Baron JC: Functional neuroimaging studies of
total motor recovery after stroke in adults: A review. *Stroke*
2003;34:1553–66
and finger tracking training in subjects with chronic
75. Taub E, Miller NE, Novack TA, et al: Technique to improve
chronic motor deficit after stroke. *Arch Phys Med Rehabil*
1993;74:347–54
cortex is implicated in continuous switching between ver-
bal fluency tasks: An fMRI study with clinical implications.
*Brain* 2002;125(5):1024–38
January 2005

fMRI and Cognitive Rehabilitation in TBI


CASE REPORT

Transforaminal Cervical Blood Patch for the Treatment of Post–Dural Puncture Headache

ABSTRACT


A 40-yr-old woman received a series of three interlaminar epidural steroid injections for the treatment of axial neck pain secondary to degenerative disc disease. Immediately after her third injection, she experienced symptoms of a dural puncture–induced headache. This headache persisted on a daily basis for 3 mos, despite two epidural blood patches using an interlaminar approach, which was finally completely abated with a transforaminal blood patch. The headache was immediately relieved and remained alleviated through the follow-up interval of 1 yr. In this patient, a fluoroscopically guided transforaminal epidural blood patch proved to be more effective than the classic blind interlaminar approach in the treatment of post–dural puncture headache.

Key Words: Post–Dural Puncture Headache, Transforaminal, Interlaminar, Epidural Blood Patch

Interlaminar epidural steroid injections are frequently performed for treatment of axial or radicular spinal pain. Post–dural puncture headache (PDH) is one of the most frequently described complications resulting from an epidural injection.\(^1\)\(^3\) It occurs secondary to an accidental dural tear, resulting in cerebrospinal fluid (CSF) leakage.\(^2\)\(^4\)\(^5\) The headache is typically transient, develops within 2–5 days, and spontaneously resolves within 2 wks.\(^1\)\(^2\) In some instances, the headache persists, necessitating medical management, including the performance of an epidural blood patch (EBP). For headaches occurring after a cervical dural puncture, the optimal location of the EBP has not been accurately described.\(^6\) In this case, we successfully employed a fluoroscopically guided transforaminal approach to inject a blood patch within the cervical epidural space after the failure of two preceding interlaminar EBPs.

CASE REPORT

A 40-yr-old female airport service attendant presented with sharp pain localized to the posterior cervical region and left shoulder, without complaints of a headache. These symptoms were acute in onset, occurring suddenly after lifting a heavy suitcase. After an initial evaluation by a primary care physician,
a magnetic resonance image of the cervical spine was ordered. This magnetic resonance image was unremarkable, except for moderate disc desiccation at the C5-C6 level. A treatment plan consisting of physical therapy and a variety of nonsteroidal oral antiinflammatory drugs failed to ameliorate her symptomatology. A referral to a pain management specialist was made. A series of three blind interlaminar epidural corticosteroid injections, performed at the C5-C6 level, was completed over a 22-day period. After the initial two injections, the intensity of her symptoms improved at least 50% on the Visual Analog Scale (3/10). Immediately after the completion of the third injection, severe position dependent occipital headaches developed. They were rated 6/10 on the Visual Analog Scale. Sitting proved to be a consistent and predictable inciting factor, whereas resting supine led to complete headache resolution. One week later, a blind cervical interlaminar EBP was performed using 9 ml of autologous blood. Because this failed to offer any symptom reduction, the next day, another blind EBP was performed, using a larger volume (13 ml) of autologous blood. Less than 50% of headache intensity reduction was realized. Five days later, the headache level intensified, surpassing that experienced before the EBPs (Visual Analog Scale, 10/10). The symptoms became constant and nonpositional. Almost simultaneously, the neck and posterior shoulder pain intensity levels returned to that experienced before the performance of any procedures. A diagnosis of myofascial pain was made, and the patient was subjected to >30 trigger-point injections. After these interventions, there was neither an improvement of her neck and shoulder pain nor her continuous unremitting daily explosive headaches. Oral agents, including carisoprodol, nefazodone, clonazepam, and a variety of narcotics, were prescribed, but no symptom relief ensued. Given the intractability of her symptoms, a referral to a neurosurgeon was made with the anticipation that a cervical fusion would be performed. Surgical intervention was dismissed by the neurosurgeon and subsequent referral to an academic spine center was made.

At initial evaluation at our center, 3 mos after the third epidural injection, she described daily severe (Visual Analog Scale, 10/10), unrelenting, position-independent bilateral occipital and frontal headaches. Rapid neck movement, Valsalva maneuvers, and bright lights were particularly provocative influences. Maintaining her cervical spine in a flexed posture while providing support to her head with both hands was the only ameliorating influence, albeit <25% of the headache intensity level. She also described posterior neck pain, midline upper back pain, bilateral suprascapular pain, and left shoulder pain. An amalgamation of medications were being used, including Oxycodone/APAP 5/325 twice a day, 0.5 mg of Clonazepam every night, and 50 mg of Nefazodone every day. On examination, she was alert and oriented to person, place, and time. There was no facial asymmetry noted, and conjugate eye movements were normal. Speech was fluent. Gait and balance were normal. Cranial nerves II–XII were intact. Neck range of motion was normal on flexion and bilateral lateral bending. Extension was diminished by 50% and right or left rotation by 20%, secondary to exacerbation of baseline pain in the posterior neck and upper back with these active or passive movements. The left shoulder examination was significant for audible and palpable crepitus during passive flexion or abduction. In addition, impingement sign was positive. Manual muscle testing of the bilateral upper and lower limbs revealed 5/5 strength, except for left shoulder abduction and forward flexion, which could not be adequately tested due to pain. Sensory examination was normal to all modalities. Muscle stretch reflexes were 2+ bilaterally and symmetrical. Hoffman test was negative bilaterally. No clonus was elicited. Fundoscopic exam was entirely normal. A list of presumptive diagnoses included tension and migraine headaches. However, the history of her complaints were temporally related to blind cervical interlaminar injections, and she had no medical history of such conditions. The myofascial dysfunction of the cervicothoracic spine was likely secondary to her underlying discogenic pain and PDPH. Management of the headache symptoms was considered the highest priority, leading to the suggestion of repeating the EBP, but using a fluoroscopically guided transforaminal approach. We hypothesized that the previous posterior interlaminar approach may have failed to direct the blood in the ideal area due to membranous raphe. Utilizing the transforaminal approach would allow us better and more accurate control over the spread of the injectate. Under real-time fluoroscopic imaging, dorsal epidural spread of contrast was observed after guided placement of a 1.5-inch, 22-gauge needle. The tip was positioned superiorly and dorsally in the foramen, and the bevel was rotated to face cephalodorsal. Six milliliters of autologous blood was instilled subsequent to outlining the dorsal epidural space.

Fifteen minutes after the delivery of 6 ml of autologous blood through the C5-C6 and disper-
sion along the dorsal epidural space, the patient reported complete and categorical abatement of her headache. At the 2-wk follow-up, the headache relief continued; however, the neck, upper back, and left shoulder pain persisted. She subsequently received three left C6 transformaminal epidural corticosteroid injections with complete resolution of the neck and upper back symptoms. Physical therapy was initiated for her left shoulder while she awaited an initial evaluation by a shoulder specialist. Telephone follow-up at 3, 6, and 12 mos revealed continued headache, neck, and upper back pain resolution.

**DISCUSSION**

The prevalence of PDPH is reported to be reaching as high as 30% after epidural anesthesia, spinal tap, and myelography.\(^1,2\) In one study on obstetric patients after delivery using epidural analgesia, its prevalence exceeded 80%.\(^7\) PDPH typically occurs within 48 hrs.\(^2\) It is usually transient, self-limiting, and usually resolves with bed rest, adequate hydration using sufficient oral fluids, and analgesics.\(^1,2\) Occasionally, it persists for more than a year.\(^8\) In rare instances, if untreated, a subdural hematoma may develop.\(^9\)

PDPH is more common in young women,\(^10\) especially those with a lower body mass index,\(^11\) and in patients with a history of headaches,\(^11\) which may be related to an enhanced sensitivity to tension applied to cerebral tissue.\(^1\) It is highly reported in the obstetric population, which may be due to high utilization of epidural anesthesia during labor. PDPH is also reported in patients undergoing surgery using spinal analgesia\(^12\) and after performing epidural or spinal injections with 22-gauge or larger spinal needles.\(^13\) Prolonged recumbency after lumbar puncture is identified to increase the risk of PDPH.\(^11\) The pathophysiology leading to the generation of a headache is believed to be the reduction of CSF volume or pressure, leading to traction and tension across the pain-sensitive dural sinuses that occurs when in the upright position.\(^1,6\) Vakharia et al.\(^14\) confirmed the CSF leak by demonstrating focal accumulation of clear or blood-stained fluid extrathecally in four of five patients with PDPH after spinal or epidural anesthesia.

Although the headache onset is usually within 48 hrs, it may be delayed for up to 12 days.\(^2\) Symptoms typically persist for <5 days in 80–85% of patients\(^1\) and are expected to disappear within 14 days.\(^15\) A CSF fistula should be considered if it lasts longer,\(^15\) which was certainly a consideration in this case. PDPH is commonly associated with neck pain and vestibular and ocular symptoms.\(^16\) Sitting upright, standing, and straining exacerbate the headache, whereas resting supine commonly relieves it.\(^1,2\) The headache is typically diagnosed by history and examination and rarely are other diagnostic investigations required. It has been reported that contrast-enhanced magnetic resonance imaging\(^17\) and superior ophthalmic veins color-Doppler flow imaging\(^18\) can reveal abnormalities in the presence of intracranial hypotension.

The clinical picture of PDPH may vary. Vilming and Kloster\(^15\) reported that 35% of their patients reported nonpostural headache. In their view, PDPH should be suspected in any patient reporting headache within 4 days of a lumbar puncture. Ferre\(^19\) reported that 7 mos after epidural anesthesia for a caesarian section, the patient continued to describe a headache that was fully relieved after the performance of an EBP. Klepstad\(^8\) reported on a 20-yr-old man admitted to the hospital with a headache of a duration of 1 yr that was accompanied by generalized fatigue and unrelated to specific postural positions. The headache was perceived as a tightness and pressure sensation within the prefrontal area. There were no associated neurologic findings or signs of increased intracranial pressure. PDPH was suspected as there was documentation of two spinal injection procedures performed to establish anesthesia during the treatment of a lower limb fracture.

The treatment of PDPH can be accomplished with conservative or interventional methods. Less aggressive methods include strict bed rest, adequate hydration, and oral analgesic agents. Caffeine is commonly used because of its ability to cause cerebral vasoconstriction without increasing CSF pressure.\(^7\) A variety of agents can be injected into the epidural space. Dextrose and saline have been used; however, the infusion of an autologous blood EBP provides the best outcomes.\(^20\)–\(^22\) Performing an EBP is the consensus treatment option for PDPH recalcitrant to noninvasive measures\(^16,23\) because success rates as high as 93% have been reported.\(^16\) However, up to 36% of patients may experience a recurrent but less severe PDPH.\(^24\) There are three theories explaining the mechanism by which an EBP can provide symptom relief. The first is that the EBP creates a blood clot adherent to the dura mater, thereby sealing a hole in the dura and preventing CSF leak.\(^25,26\) The second theory is that the injected blood leads to an increase in intraspinal pressure, effectively reducing traction on the brain and meninges.\(^6,14\) The third hypothesis is that mixing blood with the CSF leads to a rapid coagulation response, sealing the dural tear even if it is far from the blood patch injection site.\(^25\)

Gormley\(^27\) described the first successful treatment of PDPH with blood patching in 1960. Seven patients (including himself) were injected to treat headaches that developed after spinal anesthesia or
a myelogram. In each instance, relief was reported to occur within 30 mins.

The most commonly used technique for EBP involves an interlaminar injection of approximately 20 ml of blood, although the suggested volume may be as small as 2–3 ml.2,12,17 Szeinfeld et al.28 analyzed the distribution of technetium-tagged red blood cells after an EBP. They recommended that 12–15 ml of blood should be infused as close as possible to the puncture site. Such a recommendation was based on the performance of blind EBPs, as fluoroscopy was not used in their study.5,16

The transforaminal approach using fluoroscopy can result in an efficient instillation of the medicine into the dorsal or ventral epidural space, depending on needle position and bevel orientation. Indeed, this fluoroscopically guided approach is commonly used in the treatment of cervical or lumbar radicular pain.29 Huston et al.17 prospectively studied the side effects and complications of this technique. An analysis of 350 consecutive cervical and lumbar transforaminal injections identified no instance in which a dural puncture occurred. Lutz and Wisneski18 identified no dural punctures or other major complications in 50 patients who underwent lumbar transforaminal epidural injections. Botwin et al.30 reviewed complications in 322 transforaminal lumbar epidural injections done on 207 patients. They reported the complete absence of PDPH. The most common complication in their study was headaches occurring in 3.1% of patients. These headaches were transient and resolved after 24 hrs. These patients’ epidurograms were reviewed and there was no intrathecal pattern noted.

For more than a decade, we have employed this technique and have similarly observed few serious complications. More importantly, during that interval, we acquired the ability to purposefully and consistently direct medication to traverse the ventral or dorsal epidural space, depending on needle-tip placement and bevel attitude. In this instance, we placed our 22-gauge spinal needle along the superdorsal aspect of the neural foramen and oriented the bevel in a cephalodorsal direction. Contrast was then instilled to ensure proper flow along the dorsal epidural space. Thereafter, 6 ml of autologous blood was infused. The result was immediate relief of headache symptoms. Based on this case report, we recommend further scientific inquiry into the use of fluoroscopically guided transforaminal EBP. Intuitively, and as suggested by previous reports, it is a safe and accurate alternative to a blind interlaminar EBP.

CONCLUSION

This case illustrates the importance of considering PDPH as a cause of headache after interlaminar epidural injections. The diagnosis should not be ruled out simply because of failure of an EBP blood patching to relieve the headache symptoms. The diagnosis should be suspected if a meticulous history it is obtained, as there is an unambiguous temporal association with a spinal injection procedure. It must be emphasized that not all headaches classified as PDPH have the typical postural modulation of symptoms. For this patient, a fluoroscopically guided transforaminal EBP proved to be more effective than the classic blind interlaminar approach. Additional scientific studies may prove our current belief that our novel technique should be considered as the first-line interventional approach that will supplant the interlaminar method.

REFERENCES

Ultrasound-Guided Aspiration of Symptomatic Rotator Cuff Calcific Tendonitis

Grant Cooper, MD, Gregory E. Lutz, MD
Ronald S. Adler, MD

From the Department of Physical Medicine and Rehabilitation, New York-Presbyterian University Hospitals of Columbia and Cornell, New York, New York (GC); and the Departments of Physical Medicine and Rehabilitation (GEL) and Radiology (RSA), Hospital for Special Surgery, New York, New York.

All correspondence and requests for reprints should be addressed to Gregory E. Lutz, MD, Department of Physical Medicine and Rehabilitation, Hospital for Special Surgery, 535 East 70th Street, New York, NY 10021.

0894-9115/05/8401-0081/0
American Journal of Physical Medicine & Rehabilitation
Copyright © 2005 by Lippincott Williams & Wilkins
DOI: 10.1097/01.PHM.0000150821.55552.2C

Ultrasound-guided percutaneous needle puncture, aspiration, and lavage is a new, minimally invasive technique for the treatment of symptomatic rotator cuff calcific tendonitis. Approximately 10% of patients with chronic symptomatic rotator cuff calcific tendonitis are unresponsive to conservative medical management and require further intervention. In the past, either arthroscopically assisted or open-technique surgery has been the treatment of choice for these patients. Indeed, several research investigations have concluded that, in comparison with conservative therapy, surgery offers a significantly improved outcome for patients. However, surgery inevitably is accompanied by comorbidity and increased costs. Ultrasound-guided percutaneous needle puncture, aspiration, and lavage is an attractive alternative to surgery for these patients.

Ultrasound-guided percutaneous needle puncture, aspiration, and lavage easily locates and aspirates all soft calcifications in the shoulder (Figs. 1 and 2). It also fragments residual hard calcifications, which helps them to migrate to vascularized soft tissues. This fragmenting and migration accelerates the resorption process and makes it more complete. To date, five clinical studies investigating ultrasound-guided aspiration and lavage have been reported in the literature. The clinical success rates of these studies range from 60% to 74% and have follow-up times ranging from 2 wks to 1 yr. However, no controlled clinical trials have assessed the efficacy of this procedure. Future research will hopefully seek to fill this void.

In 1986, Bigliani et al. used supraspinatus outlet radiography to identify three distinct classifications of acromial morphology: type I (flat), type II (curved), and type III (hooked). The criteria of Bigliani et al. have become the most widely used criteria for acromion morphologic classification. It is a reasonable assumption that a patient’s acromion type may have a direct impact on the ability of a percutaneous needle approach to access an identified calcification in some shoulders, thus potentially affecting the efficacy of the ultrasound-guided percutaneous aspiration and lavage. Future research may investigate this potential causal relationship between Bigliani type and the efficacy of ultrasound-guided percutaneous aspiration and lavage.

REFERENCES