<table>
<thead>
<tr>
<th>Page</th>
<th>Section</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>81</td>
<td>Ethics</td>
<td>Ethical Issues in Rehabilitation Research: Conundrums for the Clinician.</td>
<td>Stanley F. Wainapel, MD, MPH</td>
</tr>
<tr>
<td>82</td>
<td>Commentary</td>
<td>The History and Moral Foundations of Human-Subject Research.</td>
<td>Jeffry Blustein, PhD</td>
</tr>
<tr>
<td>86</td>
<td>Analysis</td>
<td>Treatments to Enhance Recovery from the Vegetative and Minimally Conscious States: Ethical Issues Surrounding Efficacy Studies.</td>
<td>John Whyte, MD, PhD</td>
</tr>
<tr>
<td>93</td>
<td>Research Article</td>
<td>Functional vs. Strength Training in Disabled Elderly Outpatients.</td>
<td>David E. Krebs, DPT, PhD; Donna Moxley Scarborough, PT, MS; Chris A. McGibbon, PhD</td>
</tr>
<tr>
<td>104</td>
<td>Research Article</td>
<td>Comparisons of the Brief Form of the World Health Organization Quality of Life and Short Form-36 for Persons With Spinal Cord Injuries.</td>
<td>Mau-Roung Lin, PhD; Hei-Fen Hwang, MS; Chih-Yi Chen, MS; Wen-Ta Chiu, MD, PhD</td>
</tr>
<tr>
<td>114</td>
<td>Research Article</td>
<td>A Prospective Study to Validate an Impairment Questionnaire for Major Trauma Survivors.</td>
<td>Shawn C. Marshall, MD, MSc, FRCPC; Daryl Gray, MD, FRCSC; Keith G. Wilson, PhD, CPsych; Jean-Denis Yelle, MD, FRCSC; Paul Hébert, MD, MHSc, FRCPC; Keith O'Rourke, MBA; Maryam Jamshidi, MSc</td>
</tr>
<tr>
<td>125</td>
<td>Research Article</td>
<td>Falls and Gait Characteristics Among Older Persons with Peripheral Neuropathy.</td>
<td>Trina K. DeMott, MS, PT; James K. Richardson, MD; Sibylle B. Thies, PhD; James A. Ashton-Miller, PhD</td>
</tr>
<tr>
<td>133</td>
<td>Research Article</td>
<td>The Effects of Short-Term Exercise Intervention on Falls Self-Efficacy and the Relationship between Changes in Physical Function and Falls Self-Efficacy in Japanese Older People: A Randomized Controlled Trial.</td>
<td>Takeshi Arai, RPT, MS; Shuichi Obuchi, RPT, PhD; Yasuko Inaba, RPT, MS; Hiroshi Nagasawa, RPT, PhD; Yoshitaka Shiba, RPT, MA; Shuichiro Watanabe, MD, PhD; Ken Kimura, PhD; Motonaga Kojima, RPT, MS</td>
</tr>
<tr>
<td>142</td>
<td>Invited Review</td>
<td>Cardiovascular Disease in Spinal Cord Injury: An Overview of Prevalence, Risk, Evaluation, and Management.</td>
<td>Jonathan Myers, PhD; Matthew Lee, MS; Jenny Kiratli, PhD</td>
</tr>
</tbody>
</table>
**Case Series**

Electromyography

Jeffrey A. Strakowski, MD; William S. Pease, MD; Ernest W. Johnson, MD

**Case Report**

Spinal Cord Injury

158 Spinal Cord Infarction Secondary to Cocaine Use.
Adam L. Schreiber, DO, MA; Christopher S. Formal, MD

161 Belated Diagnosis of Congenital Myotonic Dystrophy in a Boy with Cerebral Palsy.
Se Hee Jung, MD, MS; Moon Suk Bang, MD, PhD

**Visual Vignette**

166 Epidural Corticosteroid Injections Precipitating Epidural Hematomas with Spinal Paresis.
Myron M. LaBan, MD, MMSc; Gopi Kasturi, MD; I-Ming Wang, MD
Ethical Issues in Rehabilitation Research
Conundrums for the Clinician

ABSTRACT

Key Words: Ethics, Rehabilitation, Research

Ethics have been part of our civilization since the classic writings of Aristotle. As physicians, we have sworn to uphold ethical principles enunciated by Hippocrates or Maimonides. We frequently and rather casually attest to the ethical attributes of colleagues for whom we write letters of recommendation. Those of us affiliated with hospitals have access to bioethicists when we encounter ethical dilemmas in patient care.

But how often have we considered the ethical implications of our clinical research? Given the recent spate of highly publicized episodes such as COX-2 inhibitor drugs being withdrawn from the market by pharmaceutical companies, articles published in the New England Journal of Medicine that had to be retracted because of spurious data, and the scandal over cloning research in South Korea, nonethical practices are of more than academic interest. A dearth of literature on the particular ethical issues associated with rehabilitation research stimulated the Association of Academic Physiatrists (AAP) to apply for and obtain a series of grants through the Office for Research Integrity within the Association of American Medical Colleges. The result of these grants include a previous course on rehabilitation research ethics at the AAP national meeting, a multidisciplinary course on this topic that was held at the 2006 meeting of the American Congress of Rehabilitation Medicine, and a series of articles, two of which can be found in this issue of the American Journal of Physical Medicine and Rehabilitation.

Topics to be covered include placebos in research, informed consent issues in patients lacking full decisional capacity, statistical methods, subject selection, and mentorship issues. As in the article by Dr. Whyte in this issue, we will try to focus each subject by using real-life examples that illustrate the problems and challenges faced by researchers who strive for the ideal of clinical equipoise in their work.
The History and Moral Foundations of Human-Subject Research


Key Words: Ethics, History, Rehabilitation, Research

The history of biomedical research involving human subjects is filled with instances of ethically questionable or blatantly unethical research. The most gruesome of these were carried out by the Japanese in Manchuria in the 1930s and 1940s and by German physicians in Nazi concentration camps in World War II. The pseudoscientific medical experiments of the Nazi doctors were showcased at the so-called Doctors Trial at Nuremberg, Germany, beginning in 1946. With the aid of documentary evidence and witness testimony, prosecutors exposed the cruel and inhumane procedures to which concentration camp prisoners had been subjected in the name of medical research. Twenty-three of these so-called physicians were put on trial; 16 were convicted of war crimes, and seven were sentenced to death.

The judges at the trial enunciated a set of principles called the Nuremberg Code in an effort to clarify the basis of the tribunal’s condemnation of the Nazi experiments. The Nuremberg Code, a milestone in protecting the rights of research subjects, makes explicit the ethical requirements for acceptable research and imposes extremely stringent ethical constraints on how research is to be conducted. In unambiguous and unqualified language, the Nuremberg Code begins by noting the critical importance of the voluntary consent of the subject:

The voluntary consent of the human subject is absolutely essential. This means that the person involved should have legal capacity to give consent; should be so situated as to be able to exercise free power of choice, without the intervention of any element of force, fraud, deceit, duress, overreaching, or other ulterior form of constraint or coercion; and should have sufficient knowledge and comprehension of the elements of the subject matter involved as to enable him to make an understanding and enlightened decision.1

There were also ethical abuses of human-subjects research in the United States in the aftermath of the war, although these cannot be compared in intent or outcome with the horrific and barbaric practices of the Nazi doctors. Many researchers here thought the Nuremberg Code applied only to “barbarians” and not to “civilized physician–investigators” such as themselves,2 so human-subjects research was not as strongly influenced by the principles of the Nuremberg Code as it ought to have been. Among the best known of these abuses was the research conducted at the Jewish Chronic Disease Hospital, in which foreign cancer cells were injected into the skin of 22 debilitated residents without...
informing them of the nature of the injections; the Willowbrook hepatitis studies, in which institutionalized children were intentionally injected with hepatitis in the search for an effective vaccine; and, perhaps most infamous of all, the Tuskegee syphilis study, which actually began in 1932 and continued for 40 years. In this study, to determine the effects of untreated syphilis, itinerant black farm workers from Alabama were lied to about their condition and were not given standard therapy for their disease. It was not until 1997 that President William Clinton apologized on behalf of the United States to a handful of survivors and their relatives.

By the time the details of the Tuskegee study came to light in 1972, other developments were underway that led to a closer examination of the prevailing norms of American medicine. In 1974, Congress established the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research with the following mandate: “Identify the ethical principles which should underlie the conduct of biomedical and behavioral research with human subjects and develop guidelines that should be followed in such research.” The commissioners were aware of other documents relating to the ethical conduct of research, including the Nuremberg Code and the Declaration of Helsinki, adopted by the World Medical Association in 1962, but they took as their charge to explore more deeply than before the ethical foundations of human-subjects research. The result of their lengthy deliberations was the Belmont Report, completed in 1978. This seminal document has played a key role in shaping contemporary ethical standards for the conduct of human-subjects research. The principles it proposes were also crafted into regulatory language and became public law governing the research activities of federally funded scientists.

The articulation of ethical standards for the conduct of research on human subjects is necessary, but it is not sufficient to protect the subjects’ rights. The standards also need to be backed up by effective regulatory controls. There is now mandatory oversight of most research with human subjects conducted in this country, including a review of the consent documents that are presented to prospective subjects. It is likely that these reviews have improved the quality of the research, reduced risks to subjects, and increased the amount of information given to individuals who are considering whether to become subjects. However, in the light of some recent well-publicized cases of failures to adequately protect research subjects, such as the death of Jesse Gelsinger in a gene-transfer study at the University of Pennsylvania in 1999, and the death of a healthy young woman, Ellen Roche, in a research study at Johns Hopkins University’s School of Medicine in 2001, new questions have been raised about the effectiveness of the existing oversight system.

THE BELMONT REPORT: FUNDAMENTAL MORAL PRINCIPLES

Preliminaries

The report begins by distinguishing the practice of clinical medicine from clinical research. The difference lies in the ends of the activities: in treatment, the goal is to enhance the well-being of an individual patient; in research, the primary goal is to generate generalizable knowledge. These goals can support and enhance each other, but they can also diverge.

The report also set out a definition of experimentation, distinguishing it from more colloquial uses of the term. An experiment formulates an intellectual strategy to guide interventions in a way that holds out the prospect of producing generalizable knowledge that can be validated. Research is usually described in a formal protocol that sets out the objectives of the study and the methodology to be employed to achieve them. Research, it is important to add, usually puts subjects at some degree of risk, even if only minor. The risk may be that of being randomized to an ineffective arm (or placebo) of the study or of experiencing more serious harm. Subjects become means to an end, and they may or may not experience any therapeutic benefit from their participation in the study. For this reason, special protections for research subjects are required. All these points apply to medical rehabilitation research in particular: the chief objective is to gain generalizable knowledge that can benefit future patients; there are risks, even if minor, associated with participating in the trial, including the risk of not receiving the best treatment; and special precautions must be taken to ensure that the risks are minimized.

THE THREE PRINCIPLES

Respect for Persons

This principle departs from the utilitarian outlook that governed so much research in the period after World War II. Respect for persons is a requirement rooted in Kantian philosophy and designed to protect the individual from being used to promote the greater good of the many. The principle incorporates two convictions: individuals should be treated as autonomous agents, and persons with diminished autonomy are entitled to special protection. The principle also underlies the requirement that subjects provide voluntary and informed consent before being enrolled in a research study. As part of the informed consent process, subjects must be assured that they are under no obligation...
to participate in the research and that if they agree to participate, they have the right to withdraw without being penalized for doing so.

**Beneficence (Including Nonmaleficence)**

This principle requires researchers to minimize possible harms to those absolutely necessary to achieve the research objective and to maximize possible benefits for the research subjects. It governs the conduct of individual research studies as well as scientific research in general. To follow this principle, it is necessary to do an assessment of the risks and benefits for the subjects of participating in the research. Risks may be categorized as either minimal or greater than minimal; research may hold out the prospect of directly benefiting the individual research subjects as well as others, or only others. Risks as well as potential benefits should be discussed as part of the informed consent process and should be set out in an informed consent document.

**Justice**

Aristotle gave the principle of justice a formal definition: treat like cases similarly and unlike cases dissimilarly. This means that similarities meriting similar treatment and dissimilarities meriting dissimilar treatment must be morally relevant. For example, beating one’s opponent in a competitive sport (assuming there has been no cheating) is a morally relevant basis for treating the victor differently from the loser; being a woman is not a morally relevant basis on which to exclude individuals of this sex from access to higher education. In short, differential treatment of individuals is sometimes justified, hence not unjust, and sometimes invidiously discriminatory. There are also different types of justice, including distributive justice, punitive justice, and compensatory justice. The basic idea behind the notion of distributive justice, the type of justice discussed in the Belmont Report, is that benefits and burdens should be allocated fairly and on grounds that are not morally arbitrary, so that there is some appropriate balance between benefits and burdens for all members of a particular population.

The Belmont Report does not present a theory of justice in any detail, but it does make claims about just and unjust practices. For example, imposing burdens on a particular population while allowing benefits to flow only to another population is unjust. (This point has particular relevance in connection with research conducted in developing countries.) With respect to individual research studies, subjects should not be chosen merely because they are available, or compromised, or manipulable; and in research generally, subject selection should not reflect social, racial, sexual, and cultural biases. For example, even if the researcher is not acting unfairly in selecting particular subjects, there may be larger socioeconomic and racial disparities in how subjects are selected that preclude some groups from receiving the benefits of research. It is also a matter of justice that there be an order of preference in the selection of classes of subjects (with adults coming before children) and that some categories of persons (such as the institutionalized, demented elderly) be used, if at all, only under carefully delineated conditions.

**VULNERABLE POPULATIONS**

There are categories of human subjects who generally cannot provide legally and ethically adequate informed consent or whose capacity to do so is questionable. These groups include children, adults who lack decision-making capacity, prisoners, and fetuses and embryos. The issues raised by involving children in research are complicated by the fact that there are different subgroups included under this general heading. Neonates and very young children clearly cannot participate in decisions about research, but as children grow older, they are increasingly able to understand the nature of the study and to express an opinion about it. Some of the same problems that arise in pediatric research, and others as well, are raised by research on adults who lack the capacity to provide informed consent. Although research on such individuals is necessary to develop effective treatments for individuals affected with these conditions, it is uncertain who has the authority to make decisions on their behalf and to enroll them in clinical trials. Prisoners are the classic captive population and are an especially vulnerable class of potential research subjects. Historically, they have been exploited by researchers who have viewed them as a ready source of research subjects. In all these cases, the challenge is to protect the rights and interests of research subjects without unduly restricting their access to potentially beneficial research. Finally, in our society, which is polarized by the issue of abortion, research on fetuses and embryos is especially controversial.

**THE AFTERMATH OF BELMONT**

The Belmont Report emphasizes protecting human subjects of research from harm, coercion, and nonconsensual experimentation. It is a protectoristic document, and, given the history out of which it grew, it is unsurprising that it should be so. The Belmont principles continue to exert considerable influence on the practice of research, but there has also been a significant new direction in our thinking about research, largely because of the advent of the AIDS epidemic. Central to this is a reconsideration of what justice requires in re-
search. Whereas the Belmont Report considered justice to be largely a matter of protecting the vulnerable from exploitation and manipulation, AIDS activists argued for broader access to ongoing clinical trials for patients who were willing to trade off potentially increased and certainly known risks for the possibility, however small, of benefit. Patients who were ineligible for trials because they did not fit entry criteria or who did not want to be randomized into one of the study arms wanted the option of taking investigational drugs from the trial protocol. Moreover, groups disproportionately affected by HIV/AIDS—prisoners, drug users, and women (many minority)—were excluded from trials, either because of strict protocol-entry criteria or because of lack of access to the physicians and healthcare institutions that control research. In short, broadened access to research, not protection from it, was demanded as a matter of social justice.

The movement for greater access to research for AIDS patients has largely run its course, as the activists who launched it have revised their attitudes toward alternative pathways of access to non-validated medications and have become more critical of their earlier position. Nonetheless, the shift in emphasis—from exclusion to inclusion—has had an impact throughout the research community. As a reflection of this new orientation, significant changes in National Institutes of Health and U.S. Food and Drug Administration policies have led to greater inclusion of women and children in research studies.

The Belmont principles continue to be relevant. Indeed, they may be more relevant than ever, as we have entered an era of greater protectionism in research, signaled by the following developments: third-party monitoring of consent and study procedures (data safety and monitoring boards and committees); disclosure of financial arrangements or other potential conflicts of interest; required training of investigators in research ethics and research regulation; and independent review of the decision-making capacity of potential subjects.

The research enterprise in this country has undergone significant transformation during the past several decades. There have been marked changes in how research is funded, who conducts it, and where it occurs. Once, clinical research was an enterprise primarily administered by academic researchers driven by a desire for knowledge, fame, or career advancement. Today, it is a multibillion-dollar industry, with hundreds of testing and drug companies working with thousands of private doctors. Since 1990, the number of private doctors in research has almost tripled. The industry treats research agreements with private doctors as corporate secrets and contractually forbids doctors to disclose them. This makes adequate oversight of enrollment procedures and the conduct of research extremely difficult.

Changes in how research is conducted during the past several decades have led to substantial financial entanglements between researchers, including those at academic medical centers, and for-profit companies. This development has generated increasing concern about the impact of conflicts of interest on the way that research is conducted. Conflicts of interest adversely affect the informed consent process and erode public trust in and support for the research endeavor. They are commonly addressed on an institutional level by requiring disclosure of significant financial relationships between the investigators and the private companies that sponsor their work.

A new layer of regulation has been added by the HIPAA requirements relating to human-subjects research. HIPAA’s primary research requirement governs when authorization for use and disclosure of personal health information is required from a patient and what form that authorization may take. The intent of HIPAA is not to stifle research but to provide additional safeguards for the privacy and security of health information.7

REFERENCES
7. Amatayakul M; American Health Information Management Association: Another layer of regulations: research under HIPAA. J AHIMA 2003;74:16A-16D
Treatments to Enhance Recovery from the Vegetative and Minimally Conscious States
Ethical Issues Surrounding Efficacy Studies

ABSTRACT


Randomized double-blind placebo-controlled trials have been argued to provide the strongest test of efficacy and, as such, are important tools for advancing the evidence base supporting rehabilitation treatment. However, such trials present difficult ethical issues, because one group, by definition, receives no treatment for the condition being studied. In the case of an experimental treatment that is available only within a research protocol, a 50% chance of receiving the desired treatment may be sufficient to motivate enrollment. However, many rehabilitation treatments that need further study are available outside of research protocols and are perceived as low risk, making the advantages of research participation less clear and the task of weighing the pros and cons of research participation more difficult. In this article, we discuss a placebo-controlled trial currently underway in which this issue is combined with a number of other complicating factors, such as the inability of study participants to provide their own informed consent, and the catastrophic nature of the disability under study. We examine whether other research designs could successfully answer efficacy questions in this area, and we discuss the ethical and psychosocial issues involved in planning the trial and seeking enrollment.

Key Words: Brain Injuries, Clinical Trials, Equipoise, Informed Consent

Severe brain injuries are typically associated with loss of consciousness. Although many survivors recover consciousness within hours or days, prolonged or even permanent unconsciousness may result for a minority of individuals.1 Within a few weeks of injury, true coma generally evolves into the vegetative state (VS), or individuals may transition into the minimally conscious state (MCS). After traumatic injuries, recovery of consciousness may occur after surprisingly long intervals.2 Generally, however, the longer an individual remains unconscious, the lower
the chance of regaining consciousness and the greater the chance of severe functional deficits if such recovery occurs.

Prolonged disorders of consciousness are enormously costly in economic terms, and the emotional costs to caregivers are also staggering. Caregivers are often caught between hope, which drives them to continue to seek aggressive and sometimes unproven treatments that might still restore their loved one’s function, and resignation, which might allow them to redirect emotional resources and time to other aspects of life. Thus, treatments that could hasten return of consciousness and/or raise the ultimate level of functional recovery from these devastating injuries would have enormous positive impact.

During the last two decades, many different treatments have been advocated to enhance recovery from the VS and MCS, including intensive multimodality sensory stimulation, deep-brain electrical stimulation, hyperbaric oxygen treatment, and a number of psychoactive medications such as amitriptyline, bromocriptine, and amantadine HCl. Unfortunately, despite the existence of certain theoretical arguments to support these treatments, their promise in animal models, or their suggestive effects in uncontrolled case series or small controlled trials, none of them has been studied with sufficient rigor to produce class I, or even class II, evidence of its efficacy. Definitive treatment of these conditions, unfortunately, remains virtually unchanged from its status 20 yrs ago: there are no proven interventions than can enhance recovery from VS or MCS.

### Obstacles to Rigorous Efficacy Research in the VS and MCS

Several factors limit the development of rigorous data to validate the efficacy of treatments to enhance recovery from the VS and MCS. Foremost among these is the great variability seen in spontaneous recovery. Because of this variability in spontaneous recovery, large samples are needed to conduct treatment studies that can provide definitive results. Yet, few facilities have access to large numbers of individuals in VS or MCS. To conduct definitive treatment research in this area, multicenter treatment and research systems are required, but these are challenging to organize.

The ready availability of unproven treatment interventions for off-label use is an additional obstacle.

### Sensory stimulation, hyperbaric oxygen, and experimental use of drugs that are already marketed for other indications can all take place easily outside of formal research protocols. In contrast, the only way to receive invasive procedures, such as deep-brain stimulation to enhance recovery of consciousness, is to enter a research trial. In the absence of treatments of proven value, it has become common clinical practice in many institutions to attempt one or more of these treatments on patients who fail to improve promptly in their level of consciousness. In turn, clinicians and family members may come to see these treatment attempts as the current standard of care and may be reluctant to forego them. Also, because most of these treatments are viewed as relatively safe, it may be perceived that there is nothing to lose in trying them on each patient who fails to improve. In contrast, many treatments that are newly developed for a specific condition are only available to participants in clinical trials. Moreover, these treatments often have far less safety data available so that it is easier to consider their potential risks and benefits as being in balance.

### Background of the Current Study

The Consciousness Consortium, a research network composed of eight facilities in the United States and Germany, was formed by a group of clinical investigators who treat reasonable volumes of patients in VS and MCS and who have an interest in advancing research in this area. The research network conducted a longitudinal observational study of recovery from traumatic VS and MCS in their facilities, and this study was published in 2005. This study identified several factors that were predictive of short-term recovery of consciousness among individuals who were between 4 and 16 wks after injury at the time of study entry. It also provided suggestive evidence that amantadine HCl might enhance recovery, whereas dantrolene sodium might retard it. However, the drug effects were seen on only one study outcome (Disability Rating Scale [DRS] score at 16 wks after injury, but not time until commands were followed), and although the rate of recovery during amantadine treatment was greater than before its initiation, the difference in rates was not statistically significant. Moreover, the observational nature of the study could not rule out a number of potential confounds, including the possibility that the drug was prescribed more often to individuals who were beginning to show signs of recovery.

On the basis of the findings of this longitudinal observational study, the research network decided to undertake a multicenter randomized placebo-controlled trial of amantadine HCl to include an additional clinical site and a data-coordinating center. They received a grant from the National Institute on Disability and Rehabilitation Research...
to conduct this trial. The grant was awarded to Joseph Giacino, PhD, principal investigator, at JFK-Johnson Rehabilitation Center, with John Whyte MD, PhD at Moss Rehabilitation Research Institute as coprincipal investigator. At the time of this writing, they have enrolled approximately one third of the projected sample of 182 participants. Below, we describe some aspects of the inclusion and exclusion criteria for this trial and the study protocol, and then we focus on some of the difficult ethical challenges that the planning for this study has entailed.

**Study Design**

Participants in this study are individuals in the VS or MCS as defined by their scores on the DRS and Coma Recovery Scale–Revised (CRS-R). Only individuals with traumatic injuries are eligible, because of the large difference in prognosis between traumatic and nontraumatic VS and MCS. Potential participants must meet the DRS/CRS-R eligibility criteria between 4 and 16 wks after injury and must be free of uncontrolled seizures. If they are on psychoactive drugs other than anticonvulsants at the time of potential enrollment, they must be weaned from these drugs and must continue to meet the DRS/CRS-R scoring criteria at the time of enrollment.

Caregivers of eligible participants provide informed consent through a process of substituted judgment. Once enrolled, participants are stratified based on level of consciousness (VS vs. MCS) and time postinjury (4–10 wks vs. 10–16 wks) and are randomized to either amantadine or placebo within each stratum. The dose is 100 mg twice a day for the first 2 wks. If no improvement in DRS scores occurs during that time, the dose is increased to 150 mg twice a day for a week; and if there is no improvement, the dose is increased to 200 mg twice a day for the fourth week. The primary outcome is the DRS score after 4 wks of treatment. At this point, the study drug is tapered and a final outcome score is measured 2 wks later to assess whether any group difference in 4-wk outcome reverses in the absence of continued treatment.

During the 6-wk study interval, every effort is made to avoid the use of other psychoactive drugs that might confound the results. Common medical problems seen in TBI rehabilitation have been identified along with a set of recommended treatments in descending order of confounding. Caregivers consent to having other treatments constrained in this way, but treating physicians are free to deviate from the recommended order, with those deviations recorded as protocol violations.

**Ethical Issues**

This study raises difficult ethical issues because of the serious nature of the conditions being treated and the use of a placebo. Because of our own concerns about these issues, as well as concerns about caregivers’ willingness to enroll their family members and clinicians’ willingness to support this study, we conducted a number of discussions with groups of clinicians and individual interviews with family members of former patients to understand their concerns. Thus, the primary goal of these discussions was to understand the factors that clinicians and caregivers weigh in making such an emotionally charged decision about research participation, and to understand how they interpret the current state of knowledge and study risks and benefits, rather than to determine whether such a study was ethical.

Focus groups were held at four of the participating institutions, engaging a total of 36 experienced rehabilitation clinicians of varying disciplines. Eight individual interviews at two of the sites were also conducted with caregivers of patients who would have been eligible for the trial had enrollment already begun. In both the focus groups and the individual interviews, a standardized brief presentation of the study background and design were provided, and then participants were asked to respond to several open-ended questions, as shown in Table 1. Their responses were summarized in written notes and were reviewed for common concerns and themes, in the hope that we could develop informational strategies to effectively address these concerns. The themes that emerged from these meetings are discussed below.

**Why Does This Need to be a Placebo-Controlled Study?**

Many clinicians, particularly those who are accustomed to a considerable amount of off-label use of psychoactive drugs, believe that these drugs are helpful to patients or, at the very least, safe, and that, given the dire nature of a prolonged state of unconsciousness, these drugs should be given to

<table>
<thead>
<tr>
<th>TABLE 1 Discussion questions for clinicians and caregivers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. What are your general thoughts about the study?</td>
</tr>
<tr>
<td>2. How would you feel about your family member/your patient being included in the study?</td>
</tr>
<tr>
<td>3. What concerns, if any, do you have about this project?</td>
</tr>
<tr>
<td>4. How should possible participants be approached to obtain consent for participation in this project?</td>
</tr>
</tbody>
</table>
all patients who are not undergoing brisk recovery. The fact that an individual patient might be denied an opportunity to improve must be balanced against the ongoing uncertainty about which drugs to give to whom, under which circumstances.

But could the question of treatment efficacy be answered by some other treatment design? Our answer to this was no, because the variability in spontaneous recovery makes judgment of a cause-effect relationship in pre-post designs extremely challenging. In addition, if amantadine causes an irreversible improvement in recovery, then one can’t use a crossover design, because such designs assume reversibility of the treatment effect. Even without an irreversible increase in recovery from the drug, the fact that the pace of spontaneous recovery decelerates over time makes crossover designs highly problematic in this context.

The ethics of placebo-controlled designs are predicated on the concept of clinical equipoise—that is, a real uncertainty about whether a treatment does or does not offer benefit. The concept of clinical equipoise itself, however, is somewhat simplistic. Clearly, it is unethical to conduct placebo-controlled research when one knows that an effective treatment for a serious condition is available. But what does it mean to “know” something? To know it with great confidence from several large placebo-controlled trials? To know it from some less rigorous studies that suggest a treatment benefit? To know it from one’s own clinical experience? Many knowledgeable researchers would maintain that a given treatment has not been shown to be effective with class I evidence, yet they would feel a strong desire to receive that same treatment if they were faced by a dire illness. Although the ethical requirement for equipoise was originally cast as the investigator’s requirement, later writers have suggested that the field needs to be uncertain of the answer, rather than requiring that each investigator have a neutral view of the evidence. Indeed, this seems to be a critical conceptual shift to allow medical research to proceed, because almost no treatments proceed from being completely unproven to being rigorously proven in a single study. Thus, certainty builds (and equipoise erodes) gradually with the accumulation of evidence.

There is no getting away from the fact that when a drug seems promising on the basis of basic science research, theory, and case studies, conducting definitive treatment studies involves potentially denying the placebo patients a treatment that seems promising in hopes of turning “seems promising” into “definitely works.” The ethics of this, in turn, are related to the question of how often those treatments that seem promising turn out to definitely work or be ineffective. The medical literature is replete with examples of such treatments—treatments that were thought to be effective, but that turned out, in well-conducted randomized trials, to be ineffective. Two recent examples include the case in which arthroscopic lavage or debridement of an arthritic knee joint provided no benefits over sham treatment despite the strong opinion of many orthopedists that patients improved after the procedure, and the case in which vitamin A supplements, thought to have cancer-prevention effects, were actually found to increase the risk of lung cancer. But, presumably, if small or uncontrolled studies that showed a treatment to be promising were totally irrelevant to the ultimate proof of efficacy, we wouldn’t use pilot studies of this nature to decide which treatments merited more rigorous research.

Patients Will Miss the Window of Opportunity to Improve

Many clinicians and caregivers believe that treatments that might enhance recovery need to be given relatively soon after injury. Implicit in this belief is the idea that a treatment that is effective when given early will be less effective when given later. There is little evidence to support this concern. Indeed, in the absence of rigorous evidence that a treatment aids recovery at all, one clearly can’t have definitive evidence that the treatment is differentially effective depending on the time of administration. Moreover, there are a number of instances where late intervention has been shown to be effective for some rehabilitation conditions. The origin of this belief may be that any treatment given early is more likely to get credit for the brisker pace of spontaneous recovery early after injury.

There is, however, one aspect of this concern that may be realistic in today’s healthcare environment: early after injury, the patient is likely to be in a relatively intensive treatment environment that is oriented toward clinical improvement. It may be that restorative treatments would be less likely to be implemented once the patient were living at home or in a skilled nursing facility with less intensive clinical monitoring.

Should a Caregiver of an Incompetent Patient Make Decisions of This Type?

Views vary on whether a caregiver should exercise substituted judgment to consent for research on an incompetent patient. On the one hand, there is considerable opportunity for abuse, given the vulnerability of patients in the VS and MCS. On the other hand, no treatment research on these conditions could take place if one required patients to provide their own consent. Moreover, given the grave prognosis surrounding these states, it seems likely that many individuals in this
state, if they could offer an opinion, would wish to attempt experimental treatments. We can’t know, however, whether these patients would wish to participate in placebo-controlled research or whether they would merely wish to receive these treatments off label.

Do We Have to Deny Patients Other Psychoactive Treatments?

Aside from the possibility of placebo assignment, the other constraint imposed by this trial is the attempt to minimize the use of other psychoactive drugs during the data-collection interval (6 wks). From a scientific perspective, this is crucial, because free use of a wide range of other psychoactive drugs, some with overlapping agonist or antagonist actions, would surely undermine the analysis of the study endpoints. Indeed, choice of the 6-wk study interval is already a compromise related to this issue. The question that clinicians and caregivers would most like answered is whether this or any other treatment could improve the long-term recovery from the VS and MCS. However, because administration of other psychoactive medications and treatments at any point between the experimental treatment and the final outcome assessment could confound the results, conducting a 1-yr outcome study (for example) would require minimizing other experimental interventions for a year. It seems that neither caregivers nor clinicians would support such a design and that, even if caregivers consented to such a trial, there would almost certainly be differential attrition depending on whether study patients were recovering well (and, hence, were willing to forego other treatments) or poorly (in which case caregivers would be anxious to try some other intervention).

Although restrictions on other psychoactive medications are an obstacle, we do not believe they are unethical, for several reasons. First, just as there is no class I evidence for the efficacy of amantadine in promoting recovery, there is similarly no class I evidence for a range of other psychoactive treatments for problems such as agitation, various cognitive impairments, etc. Moreover, for most of the conditions seen in TBI, there is a nonpsychoactive drug that is used to treat the condition, and there is no persuasive evidence that the nonpsychoactive drug treats the condition less well than a psychoactive alternative. Thus, treating clinicians are asked to implement treatments for a coexisting condition in order of preference (based on risk of confounding). They are not compelled to do so, and patients who are treated with “prohibited medications” are still retained in the analysis because of the intention-to-treat analytic design. The lack of evidence of superiority of one treatment over another is discussed in the informed consent document, and caregivers understand that the treating physician will likely follow the prescribed order unless there is a strong reason not to.

Response to Caregivers’ and Clinicians’ Concerns

Having spent a number of years seeking useful experimental designs and discussing the ethical implications of different designs, we believed that the basic study design was ethical. We had already anticipated the feedback that prolonged periods of placebo treatment would be untenable, which had led us to propose a relatively short-term trial. However, the feedback we received helped us to develop educational materials to support the informed consent process and to more sharply characterize the risks and benefits of the trial in the consent document. In particular, it became clear that helping caregivers understand that the variability in spontaneous recovery essentially precludes judging a drug’s effect on the individual was essential, because many individuals are used to judging a treatment’s benefit by observing changes that follow it. Second, it was clearly important to note that medical research is replete with examples of promising treatments that have turned out to be useless or harmful. However, an honest informed-consent discussion had to acknowledge that two alternatives to participating in this research were to seek access to the very same drug outside the research protocol (for families that were desperately seeking a drug intervention) or to avoid drug treatments altogether (for families that viewed interference with the natural healing process as a high-risk proposition). Thus, in the end, caregivers needed some level of altruistic concern for effective treatment of future patients, to motivate participation in this study.

Relevant Precedents

Two previous research programs in other clinical areas present ethical and research design issues that are useful to consider for their applicability to brain-injury research. The first concerns the initial research on AZT (now zidovudine) in AIDS. The initial treatment trials of AZT did not face the problem of off-label use because AZT was not available except within a research protocol. However, the rapidly lethal nature of AIDS at that time caused very strong pressure to adopt alternative experimental designs rather than placebo-controlled trials. In the end, it became clear that AZT would not be approved for sale without a placebo-controlled trial, but the concern about the ethics of placebo use in this condition led to the design of rather short outcome-assessment intervals. When the drug was shown to reduce the mortality rate during this short interval, the drug was approved and made widely available. It was only as its use
became widespread and clinicians noted that initial improvements were often followed soon after by deterioration and death that it became clear that AZT did not have lasting efficacy. At that time, there was widespread public argument about whether AZT was an effective life-saving treatment or a death-accelerating poison. That argument no longer rages because it has become clear that subsequent generations of HIV treatments have relatively prolonged efficacy.

This example illustrates how ethical pressures to avoid or minimize the placebo treatment interval can, in some cases, delay the development of a clear picture of the ultimate costs and benefits of the treatment. If amantadine is shown to enhance recovery from VS and MCS during a 4-wk treatment interval, and if that recovery is shown to persist during the 2-wk washout interval, the next logical questions would be, “Would a longer treatment further enhance recovery?” and “Does this 4-wk treatment cycle affect long-term recovery?” The first of these questions might be answered by a future study in which two groups receive amantadine for different lengths of time. Ironically, a positive result from the current study might make it nearly impossible to answer the second question, because the most efficient study design for answering it would require one group to receive placebo for a prolonged period—something that would violate clinical equipoise.

The second example to consider pertains to partial lung volume–reduction surgery, which is used to treat emphysema. This treatment, in which hyperinflated emphysematous portions of lung are resected to make room for expansion of healthier lung tissue, evolved as a potential treatment, and was being sought by individuals with emphysema and paid for by their insurers. As increasing numbers of Medicare claims for this procedure mounted, the agency became concerned about the financial consequences of widespread adoption of a costly but unproven treatment. The Health Care Financing Administration implemented a policy that allowed payment for this procedure only within approved randomized controlled trials.11 This, then, provided a very large incentive for individuals to consent to participate in the trial.

Interestingly, if investigators were to provide very large financial incentives to subjects to participate in a high-risk trial, this would likely be viewed by institutional review boards as coercive, thus violating the notion of distributive justice, because the risks (and benefits) of the research might be borne disproportionately by low-income individuals. Indeed, it is unknown whether the implementation of this payment policy resulted in high-income individuals still seeking elective emphysema surgery while low-income individuals entered clinical trials.

The ethics of this approach in some sense depend on the tension between individual and community benefit. If, on the basis of poor and potentially misleading data, large numbers of patients receive costly treatments that are actually ineffective, then society’s resources are wasted and the pace of scientific development is slowed by the public’s reluctance to forego the available but unproven treatment. On the other hand, denial of insurance coverage for unproven treatments outside of clinical trials removes the individual’s freedom to weigh the existing evidence and decide whether it is strong enough to motivate him or her to undertake the treatment. So far, in the field of rehabilitation payment, payors have sought means to stop paying for experimental treatments without demonstrating much interest in supporting the clinical trials that might prove these treatments efficacious. That would seem to be the least ethical option of all, because it perpetuates the lack of firm evidence to guide public policy, while also restricting individual decision making and choice.

Summary

Definitive evidence of treatment efficacy is sorely lacking in many areas of rehabilitation. There are multiple reasons for this, but one is the relatively infrequent use of placebo-controlled parallel-group designs. Such designs are, on the one hand, perhaps the most powerful tool for clearly showing a treatment effect; on the other hand, such designs are perhaps the most ethically challenging approach to research. When used in the context of a devastating disability, and with research subjects who cannot provide their own informed consent, the ethical complexities mount. Continued discussion among caregivers, clinicians, researchers, and bioethicists will be needed to refine our notions of when such trials are warranted and to obtain public support for enrollment in them. Moreover, this ethical balancing act will ultimately be shaped by the value of the knowledge we gain from such trials and from attempts to design alternative approaches.

ACKNOWLEDGMENTS

The author would like to thank Dr. Giacino and the other coinvestigators on this trial for review of the manuscript, Monica Vaccaro, MS, and Mark Sherer, PhD for their work to understand the ethical concerns of participating clinicians and family members, and Mary Czerniak, for assistance in preparation of the manuscript.

REFERENCES

1. Whyte J, Laborde A, DiPasquale MC: Assessment and treatment of the vegetative and minimally conscious patient, in


Functional vs. Strength Training in Disabled Elderly Outpatients

ABSTRACT

Objective: To determine whether high-intensity functional training (FT) or strength training (ST) better enables impairment, disability, and functional gains among disabled community-dwelling elders.

Design: Randomized, blinded, prospective clinical trial in a large, tertiary care outpatient rehabilitation department. Fifteen elders (62–85 yrs old) referred for physical therapy with one or more impairments, including lower-limb arthritis, participated in 6 wks of FT (weekly outpatient and three to five times per week of home practice in rapid and correct execution of locomotor activities of daily living, including gait, stepping, and sit to stand) or progressive resistive ST using elastic bands with intensity, therapist contact, and home practice similar to those of FT.

Results: Both groups significantly improved their combined lower-extremity strength (hip abduction, ankle dorsiflexion, knee flexion, ankle plantarflexion, and knee extension) \((P = 0.003)\), but no statistical difference between the ST and FT group gains \((P = 0.203)\) was found. Subjects in both interventions improved their gait speed, but the FT group improved more than the ST group \((P = 0.001)\). During chair rise, the FT group improved their maximum knee torque more than the ST group \((P = 0.033)\), indicating that they employed a more controlled and efficient movement strategy.

Conclusions: These data suggest that an intensive FT intervention results in strength improvements of comparable magnitude as those attained from ST and that FT also confers greater improvements in dynamic balance control and coordination while performing daily life tasks.

Key Words: Elders, Strengthening, Functional Training, Exercises, Functional Limitations, Disability, Intervention Study
Scant data exist on the effects of functional training (FT) or strength training (ST) among older persons, and even fewer studies address elders with disabilities. The dominant paradigm in rehabilitation has been to treat the impairment (e.g., weakness) under the assumption that functional limitations will improve as well. According to Hopp, “Although training increases force-generating capacity, little is known about its effects on functional performance. Unless investigations are conducted in which different measures of functional performance are made before and following resistance training, the validity of this approach to improving the quality of life of older persons cannot be established.”

We and others have recently demonstrated that among community-dwelling elders, both strength and functional performance improves after even modest strength gains. Fiatarone et al. showed that frail, institutionalized subjects aged 90 ± 1 yrs old experienced highly significant strength gains (174 ± 31%) after an 8-wk high-resistance exercise training program. Fiatarone et al. did not report functional locomotor benefits or the real-life role changes (if any) that resulted, except that mean tandem gait velocity improved 48% from baseline measures, two subjects no longer used canes to ambulate, and one of three subjects who was initially unable to rise from a chair without arm use became able to do so after the ST. Lord et al. and Chandler et al. separately reported that 10 wks of general activity and strengthening exercise improved the functional capacity of elders. Chandler et al. specifically noted that, among 100 disabled community-dwelling elders, gait, chair rise, and stair climbing, but not balance or disability, were positively impacted by strengthening exercise. Tinetti et al. combined home-based strengthening and FT (median length 12 wks) for 104 elders and found significant locomotor activities of daily living improvements. However, no reports differentiate locomotor functional changes resulting from FT compared with conventional ST, or the compensatory mechanisms that could be engendered by impairment (organ level) or functional (whole-person level) training.

We and others suggest that late-stance ankle plantarflexion power is a critical determinant of dysfunctional gait; impaired elders have dramatically reduced ankle-power peak magnitude compared with healthy elders of the same age and stature. Furthermore, there is some evidence that decreased knee-flexion power and, in particular, increased hip-flexion power in mid to late stance phase are, in part, a compensatory response to diminished ankle-muscle power output, which seems to be only moderately related to muscle strength. Practicing lifting, gait, and other locomotor activities may confer similar benefits as high-intensity ergometry for elders. Thus, a high-intensity FT program focused on improving ankle power during locomotion should decrease the functional limitations of impaired elders.

We studied ankle-power flow resulting from 6 wks of FT (weekly outpatient and three to five times per week of home practice in rapid and correct execution of locomotor activities of daily living, including gait, stepping, and sit to stand) or progressive resistive ST using elastic bands with intensity, therapist contact, and home practice similar to those of FT. We hypothesized that both ST and FT groups would increase their muscle strength, but that those in the FT group would obtain greater functional gains.

**METHODS**

**Subjects**

Fifteen elders (62–85 yrs of age) consented to participate and signed the institution’s approved human research consent form (Table 1). Eligibility

<table>
<thead>
<tr>
<th>TABLE 1 Subject characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment Group</td>
</tr>
<tr>
<td>------------------</td>
</tr>
<tr>
<td>Strength training</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Functional training</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Total sample</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

For diagnostic classification, the presenting complaints from the patient’s “reason for physical therapy visit” were used. BMI, body mass index (kg/m²); O, orthopedic, chiefly osteoarthritis; N, neurologic, including stroke, peripheral neuropathy; A, all others, including diffuse balance problems, cardiovascular, etc. Values are means (SD) for continuous data.
criteria included age 60 yrs or older, cognitive intactness, and ability to ambulate independently for at least 15 feet. Each subject had at least one lower-limb impairment, and each had at least one functional limitation on the Short Form 36 (SF-36) nine-item physical function inventory (excluding the vigorous-activity item). All had complaints of pain and symptomatic lower-limb arthritis (Kellgren 2 or 3) for which they were referred to physical therapy. Table 1 provides demographic characteristics of the participants. Subjects were excluded because of terminal illness, progressive neurological disease, major loss of vision (legally blind), acute pain, and nonambulatory status. Subjects were recruited through weekly screening of the outpatient physical therapy appointments for subjects 60 yrs or older who received permission from their physical therapist and/or primary care physician. All 15 subjects were consecutively (without excluding any one who met eligibility criteria) entered until 15 subjects were entered; there were no dropouts. Each subject underwent detailed biomechanical analysis of locomotor activities and lower-extremity strength testing. This assessment occurred at the initial visit and 6 wks after participation in one of two exercise programs. Additional questionnaires regarding self-perceived functional ability were administered, including a modified version of the SF-36, at the initial and final visits. Laboratory staff was blind to treatment-group membership, as were patients, who were told that they had been assigned to one of two different kinds of exercise. Each patient signed an institutional review board–approved informed consent.

**Intervention**

Patients were randomly assigned via a computer-generated table to 6 wks of ST or FT. Each exercise session begins with a warm-up period (10 min) followed by an exercise period (30 min) and a 10-min cool-down/activity period. The ST intervention exercise program is a validated program for elders, has been proven to generate both strength and functional benefits, and is based on resisted proprioceptive neuromuscular facilitation exercise patterns using a series of graded resistance elastic bands prescribed by the therapist to permit 10-repetition maximum, at which point the resistance is increased to permit only 6-repetition maximum, and then it is increased again. The FT program is a novel movement-control program that has not been previously described; the program’s details are described in Table 2. This program consists of exercises simulating locomotor activities of daily living performed at three different speeds (self-selected, fast, and slow speeds) with progressive levels of difficulty. Task difficulty was graded by, for instance, having subjects hold objects during gait, increase speed and number of repetitions of each task, changing environment (including the floor surface and step height), or combining tasks to challenge balance and motor control (Table 2). Two physical therapists instructed the subjects in either FT or ST; neither therapist was involved in testing (testers were blinded to treatment group, but training therapists could not be). Both interventions were individualized to the needs of each subject in accordance with the planned algorithm to promote consistency in application of the interventions. Both exercise interventions were designed with a four-level normal progression and four additional advanced levels to obviate ceiling effects. Subjects were encouraged to exercise 3–5 days/wk, and all subjects submitted weekly exercise adherence logs, which the therapist checked at the weekly outpatient visit.

**Lower-Extremity Strength**

Bilateral lower-extremity isometric muscle-strength testing was performed immediately before the gait trials, using a handheld Compufet dynamometer (Hoggan Health Industries model # 5025.) Knee extension and hip abduction muscle-strength testing were performed with the subjects sitting. During hip abduction, the subject pressed both knees outward while abduction was blocked by the Compufet being pressed against a rigid, upright surface. Knee flexion muscle strength was tested in prone position, and plantarflexion muscle strength was tested in long sitting (sitting with the knees in full extension). Our pilot studies showed that disabled elders required these modified testing positions to prevent undue discomfort and to permit maximum muscle performance. One practice trial was performed, and the average of two recorded trials was used for data analysis; we have previously reported the reliability of these strength measures among similarly disabled elders to be \( r = 0.87. \)

**Locomotor Activities of Daily Living**

Full-body analyses were performed at the motion laboratory, including kinematics (linear and angular motions) and kinetics (forces and moments that cause these motions) during standing and locomotor activities. The instrumentation includes two Kistler piezoelectric force plates, four Selspot II optoelectronic cameras, and 64 infrared light-emitting diodes (irLEDs), attached securely to 11 body segments (right and left feet, shanks, thighs and arms, pelvis, trunk, and head). This data-acquisition system yields 6-df kinematics of the 11 body segments. Instrumentation and processing of raw kinematic data yields resolutions of \(< 1\) degree orientation and \(< 1\) mm position. All tasks were performed barefoot, with close guarding by at least two assistants to prevent falling. Data from two trials...
were collected in a given session. The tasks included gait, chair rise, and quiet standing balance tests. Below, each task is described, and the variables obtained for each activity are defined.

**Gait**

All subjects walked approximately 10 m without an assistive device, first at their preferred speed and then paced (120 bpm) by a metronome, to provide cadence-controlled between- and within-subject comparisons. Data collection began when each subject entered the viewing volume, approximately 3 m beyond the starting point. Each individual was asked to walk at their preferred pace “in as straight a line as possible as if you were taking a brisk walk through the park.” The instructions for 120-bpm metronome-paced gait were identical except that the subjects were asked to walk at a specified pace. All subjects performed one practice gait trial before data collection.

Average gait velocity is determined from change in whole-body center of gravity (CG) displacement relative to change in time during a complete gait cycle (time elapsed between consecutive ipsilateral heel strikes). Double-support duration is obtained from one of the two periods during which the body is supported by both limbs, expressed as percent gait cycle. Maximum moment arm is the Pythagorean

---

**TABLE 2** Description of the exercise intervention programs

<table>
<thead>
<tr>
<th>Warm-Up/Cool-Down Exercises*</th>
<th>Strength-Training Exercises</th>
<th>Functional Training Exercises</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sitting</td>
<td>All exercises were instructed with the goal of attaining 10 repetitions.</td>
<td>All exercise activities were performed at three speeds (self preferred, fast as possible, and as slow as possible).</td>
</tr>
<tr>
<td>1. Deep breaths (five reps)</td>
<td>Sitting</td>
<td>Sitting</td>
</tr>
<tr>
<td>2. Terminal knee extension</td>
<td>Sitting</td>
<td>1. Chair rise</td>
</tr>
<tr>
<td>with ankle dorsiflexion and</td>
<td>1. Hip flexion and resisted hip extension.</td>
<td>2. Forward reach to opposite foot</td>
</tr>
<tr>
<td>plantarflexion (five reps for each leg)</td>
<td>2. Hip flexion with external rotation and resisted hip abduction and internal rotation.</td>
<td></td>
</tr>
<tr>
<td>3. Bilateral shoulder flexion (five reps)</td>
<td>3. Ankle dorsiflexion with resisted plantarflexion.</td>
<td>3. Forward walking</td>
</tr>
<tr>
<td>4. Shoulder shrugs (six reps)</td>
<td>Sitting</td>
<td>4. Side step walking</td>
</tr>
<tr>
<td>Sitting</td>
<td>4. Toe raises</td>
<td>5. Combined forward and backward walking patterns</td>
</tr>
<tr>
<td>5. Forward lunges (five reps for each leg)</td>
<td>5. Hip flexion with resisted hip extension and abduction.</td>
<td>6. Step up/step forward and down</td>
</tr>
<tr>
<td>6. Lateral lunges (five reps for each leg)</td>
<td>6. Hip flexion and with horizontal adduction of opposite shoulder and resisted hip abduction/extension with shoulder abduction/flexion.</td>
<td>7. Marching</td>
</tr>
<tr>
<td>All exercises were instructed with the goal of attaining 10 repetitions.</td>
<td>Sitting</td>
<td>8. Stoop/squatting</td>
</tr>
<tr>
<td>8. Resisted knee flexion</td>
<td>Sitting</td>
<td>Sitting</td>
</tr>
<tr>
<td>10. Shoulder protraction and resisted retraction</td>
<td>Mode of progression</td>
<td>Mode of progression</td>
</tr>
<tr>
<td>Sitting</td>
<td>1. Four levels of difficulty. When 10 repetitions were performed (for at least 8 of 10 exercises) for 1 wk of exercise successfully, then the next level of difficulty band was introduced.</td>
<td>1. Four levels of difficulty. When 10 repetitions were performed (for at least 8 of 10 exercises) for 1 wk of exercise successfully, then the next level of difficulty band was introduced.</td>
</tr>
<tr>
<td></td>
<td>2. Depending on the specific exercise, various items from the home were used to add difficulty (e.g., plastic bottle and laundry basket).</td>
<td>2. Depending on the specific exercise, various items from the home were used to add difficulty (e.g., plastic bottle and laundry basket).</td>
</tr>
<tr>
<td></td>
<td>Advanced Levels 1–4:</td>
<td>Advanced Levels 1–4:</td>
</tr>
<tr>
<td></td>
<td>Continued progression of task-dependent difficulty and increased emphasis on combined activities, motor control, and balance (e.g., picking up laundry basket and carrying it while walking over an obstacle).</td>
<td>Continued progression of task-dependent difficulty and increased emphasis on combined activities, motor control, and balance (e.g., picking up laundry basket and carrying it while walking over an obstacle).</td>
</tr>
<tr>
<td></td>
<td>* Warm-up/cool-down exercises were identical for both functional training and strength-training exercise interventions.</td>
<td>* Warm-up/cool-down exercises were identical for both functional training and strength-training exercise interventions.</td>
</tr>
</tbody>
</table>

* Warm-up/cool-down exercises were identical for both functional training and strength-training exercise interventions.

---

were collected in a given session. The tasks included gait, chair rise, and quiet standing balance tests. Below, each task is described, and the variables obtained for each activity are defined.
Chair Rise

For each chair-rise trial, the participant was seated with the greater trochanters approximately 4 cm from the front edge of an armless, backless chair, adjusted to 100% knee height (the distance from the right medial tibial plateau to the floor). Two chair-rise trials were recorded for each subject, after performing one practice trial. Participants placed their feet 10 cm apart, at 18-degree ankle dorsiflexion. Arm and foot position constraints were used to improve the consistency of the body position during chair-rise testing.32–34 When participants kept their arms folded in a constrained position in front of their abdomen while their feet remaining in the same position throughout the chair-rise trial, the trial was considered successful and was used for analysis. Each participant was asked to “arise from the chair as you normally do” beginning on “go” after the cue “one, two, ready, go.” Chair-rise trials were performed first with eyes open then eyes closed.

Maximum trunk flexion (relative to global/room coordinates) and maximum knee torque were recorded between the start of movement and the end of chair-rise time (EOR).1,25 The start of movement is defined as the time at which anterior upper-body linear momentum is positive, and EOR is the time at which whole-body CG reaches it first peak vertical position.34 Chair-rise cycle time was calculated as the time elapsed from the start of movement through EOR. Time of lift-off, as a percentage of chair-rise cycle time, is defined as the instant the thigh segment is displaced sagittally 2 degrees.34 The maximum range of AP linear momentum was recorded between EOR and the end of data collection, approximately 5 secs.

Quiet Standing Balance Tests

Standing-still test positions were gradated to be progressively more difficult (levels 1–6), as validated by Lord et al.35 Base of support was varied by controlling foot placement into one of three positions; subjects were tested both in light (eyes open) and in darkness (eyes closed). For the feet-together position, subjects’ feet were side by side, with the malleoli 1 cm apart. For semitandem standing, the feet were also 1 cm apart, but the heel of the forward (dominant) foot was even with the toe tip of the hind foot. Foot dominance was the subjects’ swing foot when pantomime-kicking a ball. In all cases, the feet were parallel to each other and were aligned with the sagittal plane. If subjects successfully completed the semitandem standing eyes-open trials, they were asked to perform increasingly more challenging standing tasks such as semitandem standing with eyes closed, and tandem (one foot in front of the other) with eyes open and with eyes closed, until they were unable to stand without taking a protective step.35 Subjects were told to stand as still as possible during the trials. Two quiet standing trials were recorded for each subject, after performing one practice trial. The subject stood for 10 secs, after which 7 secs of data were collected. The phase-plane combined stability parameter29 was calculated to quantify standing balance control. The stability parameter is the root mean square variance of the position and velocity in horizontal (AP and mediolateral) planes.29

SF-36

Subjects completed the SF-36 in interview format. These interviews were carried out initially within 1 wk of the first test session and again after the 6-wk intervention exercise program. Within a week of completion, all subjects received (via mail) a study survey that was designed to obtain subjective reactions to the exercise programs.

Data Analysis

Descriptive statistics and repeated-measures multivariate analysis of variance with honestly significant difference–corrected paired comparisons were calculated on the outcome variables after determining the independent variables were free of multicollinearity. Although one could argue that multivariate analysis of variance on a small sample such as ours might lead to type II errors, we found statistical significance; thus, it must be concluded that the multivariate analysis of variance, combined with the conservative honestly significant difference–corrected paired comparisons, prevented potential type I errors on the multiple comparisons. The independent variable was type of intervention: ST or FT. The average of trials 1 and 2 for each muscle group was taken across both the right and left sides at both the 0- and 6-wk visits. We used an intention-to-treat analysis for overall hypothesis testing and one-tailed testing because the hypotheses are unidirectional. Pearson correlation analysis determined the relationship between right and left muscle strength and between
the subject characteristics and percent changes in strength between the 0- and 6-wk visits. A Mann–Whitney nonparametric one-tailed comparison was used for the SF-36 data analysis. SPSS version 10 (Chicago, IL) and SAS version 6.04 (Cary, NC) statistical packages were used for all analyses.

RESULTS

Subjects performed the exercise program 4.99 ± 1.07 days/wk in the ST group and 5.39 ± 1.27 days/wk in the FT group (P = 0.29). On average, the maximum exercise-level changes from baseline did not differ between the ST group (3.60 ± 1.52) and the FT group (4.38 ± 0.74; P = 0.10); each group had a four-level normal progression and four additional advanced levels, for eight total possible advancement levels.

Strength

The total sample improved their combined lower-limb strength (sum of the five muscle groups) by 19% relative to initial values (Table 3, P = 0.003). The most significant muscle-group strength gains were for hip abductors (P = 0.02) and ankle dorsiflexors (P = 0.006). There also were gains in hamstrings (P = 0.038) and ankle plantarflexors (P = 0.039) across all subjects. Though not statistically significant, there were modest gains in knee-extension strength across all subjects (P = 0.052). There were no differences between the FT and ST groups in strength gains, but, on average, the FT group improved by 25.6%, and the ST group improved 15.6% after the 6-wk intervention. The left and right baseline lower-extremity strength values strongly correlated to each other (hamstrings r = 0.88, P < 0.001; quadriceps r = 0.81, P < 0.001; P = 0.007; ankle dorsiflexion r = 0.87, P < 0.001; ankle plantarflexion r = 0.90, P < 0.001). The percent change of the combined strength variable did not correlate with age, weight, height, or BMI (r = 0.51 to −0.029, P = 0.051–0.917), thus supporting the generalizability of these results and obviating concerns about these potentially confounding variables in this sample.

Gait

Preferred Gait

All subjects significantly improved maximum and average CG velocity (P < 0.003). Overall, the FT group, compared with the ST group, had a greater improvement in maximum and average gait velocity (P = 0.024 and P = 0.023, respectively, Fig. 1). Indeed, the FT group had threefold gait-velocity improvements compared with the ST group. The FT group had significantly (P = 0.037) greater improvement (decreased percent cycle time) in double support after the intervention (FT group, −2.48% ± 2.09; ST group, −0.629% ± 1.19; Fig. 2).

Paced Gait

Three subjects (two subjects in FT and one in the ST group) were unable to walk at the ideal paced time of 120 bpm (±0.01 bpm) and were therefore eliminated from the paced-gait analysis. The 12 subjects able to perform paced gait demonstrated improved maximum moment arm and stance duration (P = 0.044 and 0.045, respectively). After the exercise intervention, the ST group showed statistically greater (P = 0.045) mean change in maximum moment arm (2.76 ± 2.64) than the FT group (0.32 ± 1.92) (Fig. 2). The ST group also statistically (P = 0.023) improved lateral linear momentum (−2.43 ± 5.94) compared with the FT group (5.11 ± 5.38) (Fig. 2).

Chair Rise

Eleven subjects were able to rise independently, without taking a step or unfolding their arms, at 100% knee height at the initial visit. Two subjects were only able to rise from a chair at 120% knee height, and two others required manual assistance to rise. Data analysis for the 11 subjects revealed a trend toward improved chair-rise (eyes closed) maximum range of AP linear momentum after the end of rise, with a decrease from 9.18 to 6.05 kg-m/sec (P = 0.045). The FT group changes in momentum did not differ from the ST pre- vs. postintervention changes (P = 0.113). The FT group used less peak knee torque than the ST group.

<table>
<thead>
<tr>
<th>Muscle Action</th>
<th>Functional Training Group</th>
<th>Strength-Training Group</th>
<th>FT–ST difference significance, P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hip abduction</td>
<td>16.4 (22.8)</td>
<td>13.7 (12.9)</td>
<td>0.40</td>
</tr>
<tr>
<td>Knee flexion</td>
<td>9.5 (15.0)</td>
<td>19.8 (41.1)</td>
<td>0.25</td>
</tr>
<tr>
<td>Knee extension</td>
<td>33.6 (49.4)</td>
<td>3.7 (30.4)</td>
<td>0.11</td>
</tr>
<tr>
<td>Plantarflexion</td>
<td>20.2 (32.5)</td>
<td>12.6 (14.9)</td>
<td>0.30</td>
</tr>
<tr>
<td>Dorsiflexion</td>
<td>48.1 (66.6)</td>
<td>28.3 (34.5)</td>
<td>0.26</td>
</tr>
</tbody>
</table>

Note that FT, on average, gain 26%, and the ST group gain an average of 16%.
group during chair rise with eyes open ($P = 0.033$) and the chair rise with eyes closed ($P = 0.031$) (Fig. 3). There were no significant differences ($P > 0.05$) in change scores for all subjects combined or between intervention groups for chair-rise cycle time and maximum trunk-flexion angle.

**Quiet Standing Balance**

Eight subjects showed improved ability in quiet standing balance, progressing one or more position level(s) of difficulty (Table 4), irrespective of group. Phase-plane analysis demonstrated an overall improvement in the CP stability parameter across the 10 subjects who performed with feet together and eyes closed during both visits ($P = 0.007$). The ST group performed better than the FT group ($P = 0.038$) in standing still, decreasing their average CP phase-plane stability parameter ($-0.22 \pm 0.31$) compared with the FT group (1.03 $\pm$ 1.54). There were no significant changes noted

**FIGURE 1** Maximum within gait-cycle (left bars) and average (right bars) preferred walking speed in the functional training (FT) and the strengthening (ST) groups: absolute changes (cm/sec) before and after exercise intervention. Larger values indicate improvement. Note that the improvement in the FT group is threefold that of the ST.

**FIGURE 2** Percent change in preferred gait-speed double-support time, paced-gait maximum moment arm, and range of lateral linear momentum between FT and ST groups. Smaller double-support time and lateral linear momentum, but larger maximum moment-arm values, respectively, indicate improvement. Note that the FT group had a threefold improvement in double-support time, but the ST group improved their maximum moment arm (CG–CP Pythagorean distance of separation) nearly fourfold that of the FT group.
in CG stability parameters during feet together with eyes closed or CG and CP stability parameters during semitandem with eyes open.

**Questionnaires**

One subject in the ST group sustained an unrelated fall halfway through the 6-wk intervention, resulting in injury of her dominant shoulder. Although she continued to participate in the intervention, we modified the exercises to use only the other arm. All subjects, with the exception of the person who sustained the shoulder injury, completed the SF-36 physical functioning subscale questionnaire initially and on completion of the exercise program. Seven subjects reported improvement in the SF-36 items. Five of these subjects were in the FT group, and two were in the ST group. The remaining seven (three in the FT and four in the ST group) showed no changes from their baseline scores. There was significant improvement among all subjects \( (P = 0.013) \) after the intervention. There was a statistically insignificant trend \( (P = 0.068) \) toward greater improvement in the FT group compared with the ST group.

**DISCUSSION**

Rehabilitation medicine in general and physical therapy in particular have long stated they treat the whole person, but, in fact, treatments are primarily directed at the impairment- or organ-system (here, the muscular system) level of dysfunction. Our data suggest that a paradigm shift may be needed: our FT group obtained both functional benefits and strength impairment improvements by progressively performing more vigorous functional activities. The ST subjects' data, although preliminary, demonstrate similar strength gains from 6 wks of a similar program, as did our prior sample of 132 community-dwelling elders, performing ST for 6 mos, \(^1\) in contrast to prior reports \(^3^7\) suggesting that only long-term exercise confers significant impairment gains. Although prior research has shown that functionally directed interventions improve activities of daily living outcomes, \(^3^8^,3^9\) we are not aware of other work directly comparing impairment-level (ST) with functional-level (FT) outcomes in disabled elders.

All subjects demonstrated improved maximum and average gait velocity during preferred gait (Fig. 1). However, the subjects in the FT group demonstrated significantly greater gait-velocity improvements, apparently because they practiced faster walking as their intervention. Subjects improved in some

![FIGURE 3 Change in normalized (height in meters · percent body weight; M-%BW) knee torque during preferred chair rise with eyes open and eyes closed between the FT and ST group. Note that more negative values indicate improvement; the FT group clearly decreased their knee stress, an important predecessor to decreased knee pain.](image)
gait parameters irrespective of the intervention group, partly because they had impairments that PT could be expected to help them improve, such as lower-limb arthritis. Because they are older, however, they also had several comorbidities for which they were referred to PT, thus enhancing the generalizability of these findings. The FT group also improved (decreased) their double-support time and increased their balance factor—the maximum moment arm—during preferred gait (Fig. 2). One might argue that the percent changes are small, but Judge38 reported similar balance and disability-score improvements from an explicit balance-training program for elders, and King39,40 reported that general physical conditioning yields more modest performance improvements, and yet both programs were judged worthy of incorporation into elders’ exercise regimes. During cadence-controlled paced gait, our ST group was significantly better than the FT group in increased moment arm and decreased lateral linear momentum. Further exploration of the functional merits of these ST outcomes, especially of the decreased lateral momentum, should be the subject of future investigations. The training effect in the FT group allowed gains in gait velocity that may result from a more efficient distribution of power during gait, as we have suggested elsewhere.41

These changes are consistent with specificity of exercise theories: training for peak performance of an activity is quite task specific. Training for peak performance, through repetition and practice, achieved better outcomes than the indirect method of strengthening muscles in a non–task-related manner.42 Those who received the ST showed effects related to the mechanics of gait pattern, increasing their velocity, decreasing their double-support time, and increasing their moment arm.

Increases in gait velocity can occur by increasing ankle plantarflexor power, essentially advancing the leg into stance phase faster and delivering more AP power to the CG.43 Using ankle power to advance the leg will result in higher knee-power absorption, which is accompanied by greater angular excursions of the ankle and knee, thus allowing the pelvis and trunk a more fluid translation. The increased walking speed, but greater angular excursions, may result in small or no changes in double-support time (cf. the −2.48% decrease for FT and −0.63% double-support decrease for ST, which could be a statistical type I error). The FT group probably walked faster by using a more efficient distribution of power, but the ST group did not walk faster by decreasing double support and increasing moment arm—rather, they walked faster because they got stronger, but without improving their gait style. We believe that these findings resulted from the ST intervention addressing only impairments and not functional limitations. Figure 4 proposes one model that may account for the FT and ST relationships to impairments and functional limitations observed in this study.

We found improved stability across all subjects during the time directly after EOR during chair rise with eyes closed. Performing chair rises with eyes closed increases the level of difficulty for the chair-rise task, because one cannot rely on visual feedback, and therefore, the patient is more dependent on information from proprioceptive and vestibular sources. The additional challenge from performing the task with eyes closed improved our ability to detect AP stability changes. The significantly decreased knee torque (Fig. 3) after treatment among the FT group suggests that these individuals are showing a learning effect of transitioning their anterior momentum into vertical mo-

---

**FIGURE 4** Proposed relationships among impairments, functional limitations, and disability, and the interventions targeted at impairments and functional limitations. Our data suggest that this model correctly accounts for the relationships depicted.

---

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.
mentum—the explicit goal of FT. Thus, the more healthy chair-rise strategy is referred to as the momentum-transfer strategy. We have previously shown (with a separate sample of impaired elders) that there is less knee and hip torque required when performing the efficient momentum-transfer movement strategy during chair rise. For our subjects with knee osteoarthritis, decreased knee torque probably translates to decreased pain. Several individuals in our study reported substantial improvements in their ability and ease in rising from a chair. More positive changes in functional capacity were expressed by those in the FT group than in the ST group. It is probable that a larger sample size will yield greater changes in chair-rise outcomes with increased statistical power.

Of interest is the finding that ST engendered better standing-still balance than FT but that FT better improved dynamic balance during gait and chair rise. Recent evidence demonstrates clearly that standing still requires primarily muscle stiffening—hardly a functional requisite for locomotor dynamics. Apparently, the ST group learned to better stiffen their muscles and thus stand still, but these benefits did not enhance the ST group’s preferred locomotor dynamics. The latter is obviously more relevant to patients and clinicians than standing still.

The strong correlation between left and right lower-extremity strength supports the common clinical contention that to decrease strength testing time and subject fatigue, one should test muscle strength only on one side. This conclusion is especially relevant to the disabled elderly patient.

Limitations

Both interventions engendered strength gains in this sample of elders. Although a longer intervention duration might yield greater strength changes for both groups, it would also result in several problems, including dropouts from unrelated morbidity acquired during the longer trial. Future investigations should also employ larger samples; the fact that statistical significance was found in this smaller sample, however, underscores the clinical importance of these data. The insignificant quadriceps-strength change and questionnaire data probably resulted from several factors. This small-sample pilot study, with limited representation of the elderly population, had greater quadriceps baseline strength than originally anticipated, and previous studies have shown that the greatest strength changes occur in elders with the lowest baseline strength. Additionally, test–retest reliability for handheld dynamometry decreases as knee strength increases. Finally, we examined only the physical functioning subscale of the SF-36; prior studies have also shown changes in mental health among older subjects after exercise. Clearly, a larger sample would enhance the statistical power of these findings.

Although it would be informative to compare FT to a more vigorous ST, such as high-intensity weight lifting, two factors may confound this comparison. For an intervention to be home based (obviously important to community-dwelling elders such as these), it cannot require the purchase of exotic or expensive equipment. In addition, these elders were referred to, and expected to obtain, physical therapy to ameliorate their lower-limb weakness and other impairments, but exposing them to the side effects of weight lifting (e.g., initial sprains and strains) may have decreased the voluntary enrollment in and subsequent adherence to this study. Moreover, the key issue is functional benefit—and these data demonstrate that to obtain a functional benefit, specific, tailored FT is more advantageous than ST.

CONCLUSIONS

These data suggest that an intensive FT intervention results in strength improvements of a magnitude comparable with those attained from ST, and that FT also confers greater improvements in dynamic balance control and coordination during daily life tasks. Intense FT can offer a direct, valuable mode of improving daily life activities and functional performance, but definitive conclusions must await a larger sample with more generalizable study populations. These data demonstrate the potential importance of intervening at the whole-person functional-performance level rather than simply addressing impairments such as weakness.

REFERENCES

Comparisons of the Brief Form of the World Health Organization Quality of Life and Short Form-36 for Persons With Spinal Cord Injuries

ABSTRACT


Objective: This study compared the psychometric performance of the brief form of the World Health Organization Quality of Life (WHOQOL-BREF) with the Short-Form 36 (SF-36) for people with traumatic spinal cord injuries in Taiwan.

Design: From a nationwide registry of traumatic spinal cord injuries, 187 subjects completed telephone interviews. Score distributions, internal consistency, intrainterviewer and interinterviewer test-retest reliabilities, convergent and known-groups validities, and the responsiveness between the WHOQOL-BREF (with an overall quality-of-life facet and four domains) and the SF-36 (with eight domains) were compared.

Results: Both the WHOQOL-BREF and the SF-36 exhibited low missing values (0.9–7.7% vs. 2.1–3.8%), very good internal consistencies (Cronbach’s alpha coefficients of 0.75–0.87 vs. 0.72–0.98), intrainterviewer reliabilities (intraclass correlation coefficients of 0.84–0.98 vs. 0.71–0.99) and responsive statistics (0.787–1.83 vs. 0–0.92), and fair interinterviewer reliabilities (intraclass correlation coefficients of 0.56–0.95 vs. 0.41–0.98), whereas the WHOQOL-BREF’s domains converged with the conceptually related domains of the SF-36. Nonetheless, compared with the SF-36, the WHOQOL-BREF had lower percentages of ceiling (0.0%–0.4% vs. 0.4%–63.8%) and floor (0.0%–1.3% vs. 0.4%–28.1%) values and better known-groups validity and responsiveness.

Conclusions: The WHOQOL-BREF is an appropriate generic health-related quality of life measure for persons with traumatic spinal cord injuries.

Key Words: Quality of Life, SF-36, Spinal Cord Injury, Taiwan, WHOQOL-BREF
Improvements in early postinjury care have minimized or eliminated many of the complications that previously accompanied new traumatic spinal cord injuries (SCIs), however, the complex long-term outcomes that people surviving with SCIs face after hospital discharge remain. As a result, the health-related quality of life (HRQL), a multidimensional construct primarily based on a person’s subjective appraisal of his/her own physical functioning, psychological functioning, social interactions, and somatic sensations, increasingly has become an important outcome measure for constructing a holistic view and determining the success of healthcare programs among persons with SCIs.

Although no specific HRQL measures for persons with SCIs have been widely accepted, generic HRQL measures such as the Short Form 36 (SF-36) and many others have been applied to this population. The generic measures can be applied to a variety of populations and allow for broad comparisons of the relative impacts of different diseases or health conditions. More recently, the World Health Organization (WHO) in 15 international field centers crossculturally developed a brief form of the World Health Organization’s Quality of Life questionnaire (i.e., the WHOQOL-BREF). The use of the instrument for persons with SCIs seems to be promising because excellent validity and reliability have been reported among a variety of populations across many countries. Nevertheless, its validation, particularly in comparison with other HRQL instruments, in persons with SCIs has not yet been examined.

Among studies on psychometric properties of HRQL measures in persons with SCIs, findings regarding the measures’ relation to some important prognostic factors, such as the level of injury, time since injury, age, and gender, are very inconsistent. Investigators have speculated that the inconsistencies are attributable to differences in HRQL measures as well as to the size or nature of the SCI samples. In other words, some HRQL measures may lack sufficient sensitivity to discriminate between persons with tetraplegia and paraplegia or even between employed and unemployed persons with SCIs. Nevertheless, empirical evidence of two or more psychometric measures being applied to the same persons with SCIs has not been provided to corroborate or reject those speculations.

For validating the WHOQOL-BREF questionnaire and understanding whether differences in performance exist between HRQL measures in persons with traumatic SCIs, this study compared the psychometric properties of the WHOQOL-BREF and the SF-36 among people with SCIs in Taiwan.

**METHODS**

**Study Subjects and Procedures**

Potential subjects were identified from a nationwide registry, which consisted of 809 traumatic SCI cases collected by the Head and Spinal Cord Research Group of the Neurological Society in Taiwan during the 4-yr period from July 1, 1992, to June 30, 1996. As adopted from Kraus et al., traumatic SCI was defined as an acute, traumatic lesion of the spinal cord resulting in any degree of sensory or motor deficit or paralysis, including injury to the nerve roots or cauda equina. Medical records of 113 hospitals in Taiwan, considered by the Head and Spinal Cord Research Group to have the ability to manage traumatic SCIs, were reviewed to develop the national registry. It included patients who were not transferred from other hospitals and who were coded with the International Classification of Disease as 806.0–806.9 and 968.0–968.9. In addition to the neurological severity of injury using the American Spinal Cord Injury Association 1982 standards, information on phone number, birth date, gender, educational level, and time and cause of injury also was recorded in the registry.

We used national identification numbers and names to search mortality data from 1992 to 2001 in the Department of Health, Executive Yuan, the Republic of China, and 64 subjects in the SCI registry were matched. Of those remaining in the registry, a phone number was recorded for 603 subjects. For the 142 subjects who did not have phone numbers recorded in the registry and those who could not be reached by their original phone numbers, we also searched the member lists of national and local SCI associations, and an additional 10 people were identified and contacted. These phone numbers were used to conduct telephone interviews, and a maximum of six telephone calls was made to each potential subject. When a subject agreed to participate in the study during the first call, we asked him/her to make available a time session of approximately 30 mins within 1 wk for our telephone interview. Interview procedures and interviewer attitudes on the telephone were standardized through participation in a 4-hr training course. In the telephone interview, the WHOQOL-BREF and SF-36 were administered; additionally, information on marital status and employment and a rating scale (RS) for self-perceived health status were also collected. Before the interview, a questionnaire was mailed to the subject to provide better comprehension of the questions to shorten the interview.

Of the 613 subjects with available phone numbers, 187 subjects were interviewed, 370 could not be reached by existing phone numbers, 40 had died, 1 was in a vegetative state, and 15 declined to...
be interviewed. A flow diagram of the study population is shown in Figure 1. Compared with the 809 SCI subjects, the 187 respondents did not significantly differ in age (45.6 vs. 42.9 yrs), gender (76% vs. 81% males), time since injury (7.8 vs. 7.4 yrs), injury cause (58% vs. 57% motor vehicle crashes), neurological severity (40% vs. 35% complete lesion) or associated injuries (43% vs. 47% being positive). This research was reviewed and approved by the Institutional Review Board of Taipei Medical University, Taiwan, and verbal consent was obtained from all participants.

**WHOQOL-BREF**

The WHOQOL Group originally developed a WHOQOL questionnaire with 100 items (WHOQOL-100) that allows a detailed assessment of 24 facets relating to quality of life. However, the long form may be too lengthy for studies in which the HRQL is only one variable of interest. Therefore, 2 items from the Overall Quality of Life and General Health facet and 1 item from each of the remaining 24 facets were selected to form the WHOQOL-BREF.\(^21\) These facets are categorized into four domains: Physical Capacity (seven items), psychological well-being (six items), social relationships (three items), and environment (eight items). The recall interval is 2 wks. All items are rated on a five-point Likert scale, and then domain scores are calculated by multiplying the mean of all facet scores included in each domain by a factor of 4, with a possible range of each raw domain score of 0~16. Where an item was missing, the mean of other items in the domain was substituted; however, when more than two items were missing from the physical, psychological, or environmental domain and where more than one item was missing from the social domain, the domain score was not calculated.\(^26\) Each domain score also is transformed onto a scale with a range of from 0 to 100, with a higher score indicating a higher quality-of-life. The Taiwan version of the WHOQOL-BREF was developed in compliance with WHO guidelines.\(^27,28\) In addition to comprising 26 items translated from the original WHOQOL-BREF, the version includes 2 additional items of local importance, i.e., Being Respected and Food Availability.\(^29,30\) Very good reliabilities (including an internal consistency of 0.70~0.77 and a test-retest reliability of 0.76~0.80) and validities (including content, criterion, discriminant, prediction, and construct) of this version have been reported.\(^31\) The two local items were excluded from this study to facilitate potential future international comparisons. To make use of all information in the WHOQOL-BREF, responses from the two items of the Overall Quality of Life and General Health facet also were calculated as a single score as with the scoring method for the four domain scores, even though this single facet score was not used by the WHOQOL group.

**SF-36**

The SF-36 is one of the most widely used HRQL instruments in the world that has been extensively validated within the Medical Outcome Survey. It consists of 36 items and is organized into 8 health domains: Physical Functioning (10 items), Role Physical (i.e., role limitations because of physical health problems, 4 items), Bodily Pain (2 items), General Health (5 items), Vitality (4 items), Social Functioning (2 items), Role Emotional (i.e., role limitations due to emotional problems, 3 items), and Mental Health (5 items).\(^32\) Another item that assesses changes in perceived health during the past year was not used in the study. The recall interval is 4 wks. The number of response
WHOQOL-BREF, and the SF-36 were tested using the convergence, correlations among the RS, the score indicates a better health state. To examine worst health and 100 being full health. A higher (i.e., a 100-point thermometer), with 0 being the their current states of health on a 0-to-100 scale to each subject, the RS asked subjects to rate Convergent Validity

Reliability

Internal Consistency and Test-Retest
Score Distribution

Distributions of the mean and median values were calculated for each domain. The percentage of respondents with missing values for each item and the distributions of minimum and maximum possible domain or facet scores (i.e., floor and ceiling values) were used to evaluate the difficulty of completion and the problematic score distribution, respectively. Furthermore, the normality of the domain/scale scores was tested using the Shapiro-Wilk method.

Known-Groups Validity

The ability of the WHOQOL-BREF and SF-36 to discriminate among subgroups with respect to six sociodemographic or medical characteristics of age, education, marital status, employment, time since injury, level of injury, and self-care ability was tested using the Mann-Whitney U test. These selected characteristics were previously reported to be significantly associated with HRQL among persons with SCIs. It was expected that subjects who were younger, were employed, had higher educational levels, had a spouse present, had lower levels of injury, and had self-care ability would have significantly higher scores compared with their counterparts.

Responsiveness

A stratified random sample by current employment status (yes vs. no) of 30 subjects, selected from those who had been employed before the SCI, were interviewed again on the phone to recall their HRQL at the time of injury and, consequently, 27 (13 employed and 14 unemployed) completed the WHOQOL-BREF and SF-36. According to Guyatt’s method, the responsive statistics of each domain of the WHOQOL-BREF and SF-36 were calculated by the difference in the mean changes in scores for that domain at the time of the study from the beginning of injury between subjects who were employed and those who were unemployed at the time of the study divided by the standard deviation of score changes in the latter group. Using Cohen’s criteria, a meaningful effect size of 0.2–0.5 was considered to be small, 0.5–0.8 to be moderate, and ≥0.8 to be large. SAS (Statistical Analysis Software; SAS Institute, Cary, NC) version 6.12 was used for all statistical analyses.
RESULTS

Among the 187 respondents, the means of age at injury and time since injury were 42.9 and 7.4 yrs, respectively. Of these subjects, 48 (25.7%) had incomplete tetraplegia, 28 (15.0%) had complete tetraplegia, 73 (39.0%) had incomplete paraplegia, and 38 (20.3%) had complete paraplegia. Furthermore, 151 (80.7%) of them were men, 23 (12.3%) had an educational level of senior high school or greater, 96 (51.3%) of junior high school, and 68 (36.4%) of elementary school or no formal education, 54 (28.9%) were single/divorced/widowed, 101 (54.0%) were unemployed, and 99 (52.9%) could care for themselves independently. Of the study subjects, 106 (56.7%) were associated with motor vehicle crashes, 62 (33.2%) with falls, and 19 (10.2%) with others.

The score distributions of the WHOQOL-BREF and SF-36 domains are shown in Table 1. The Shapiro-Wilk test for normality produced a P value of \( \leq 0.05 \) for all domain scores (i.e., these scores were not normally distributed). Nevertheless, for each WHOQOL-BREF domain score, the median was close to the mean, indicating that these score distributions were nearly symmetrical. In the Physical Functioning, Role Physical, and Role Emotional of the SF-36, differences between the mean and median values and large standard deviations were observed.

(28.1%), Role Emotional (20.1%), and Physical Functioning (12.2%) of the SF-36. The percentage of ceiling values higher than 5% included the Role Emotional (63.8%), Role Physical (54.4%), Physical Functioning (29.7%), and Social Functioning (10.9%) of the SF-36. The missing value rate varied from 0.9% to 7.7% for the WHOQOL-BREF and from 2.1% to 3.8% for the SF-36.

Cronbach’s alpha coefficients for internal consistency varied from 0.72 to 0.98 for the SF-36 and from 0.75 to 0.87 for the WHOQOL-BREF, indicating good internal consistency among these particular items within each domain. Although all the intraclass correlation coefficients of the intrainterviewer test-retest reliability for the WHOQOL-BREF and SF-36 were greater than 0.70, some coefficients of interinterviewer test-retest reliability were lower than 0.70, including the Overall Quality of Life and General Health (0.63) and Social Relationships (0.56) of the WHOQOL-BREF as well as Physical Functioning (0.67), General Health (0.41), Social Functioning (0.52), and Mental Health (0.57) of the SF-36.

As shown in Table 2, all of the Spearman’s correlation coefficients of the RS with the Overall Quality of Life and General Health facet and the four domains of the WHOQOL-BREF were greater than 0.4, and those of the RS with all of the SF-36's
domains but the Role Emotional ($\gamma = 0.32$) and the Mental Health ($\gamma = 0.36$) were also greater than 0.4. Furthermore, Spearman’s correlation coefficients of the conceptually related domains between the WHOQOL-BREF and SF-36 were underlined, and all of the coefficients of the WHOQOL-BREF’s domains with the corresponding SF-36’s were also higher than 0.4, with the exception of the WHOQOL-BREF’s Psychological Well-Being with the SF-36’s Role Emotional ($\gamma = 0.37$).

The abilities of the WHOQOL-BREF and SF-36 to discriminate between subgroups regarding six characteristics are shown in Table 3. Although all domains of the WHOQOL-BREF and SF-36 had significant discriminant ability between subgroups regarding employment and self-care ability, the discriminant ability differed between the WHOQOL-BREF and SF-36 with respect to other characteristics. For example, compared with younger subjects (≤43 yrs), older subjects had significantly greater scores on the Overall Quality of Life and General Health facet, Physical Capacity, Social Relationships, and Environment of the WHOQOL-BREF as well as in the Role Functioning, Role Physical, Bodily Pain, and General Health of the SF-36. In overall, the WHOQOL-BREF domains significantly discriminated between subgroups in terms of four to five characteristics and the SF-36 domains in terms of two to four characteristics.

The responsiveness of the WHOQOL-BREF and SF-36 regarding employment status is shown in Table 4. The effect sizes in three domains (overall Quality of Life and General Health facet, Physical Capacity, and Social Relationships) of the WHOQOL-BREF and 1 domain (Physical Functioning) of the SF-36 were large. The effect sizes of two domains (Psychological Well-Being and Environment) of the WHOQOL-BREF and two domains (Role Physical and Role Emotional) of the SF-36 were moderate, and those in two domains (Social Functioning and Mental Health) of the SF-36 were small.

**DISCUSSION**

Although both the WHOQOL-BREF’s and the SF-36’s domain scores exhibited few missing values, very good internal consistency and intrainterviewer reliabilities, and fair interinterviewer reliabilities, the constructs measured using the WHOQOL-BREF also converged with those of the RS and the SF-36; however, compared with those of the SF-36, the domain scores of the WHOQOL-BREF had lower percentages of ceiling and floor values, greater efficiency in discriminating subgroups with respect to the selected characteristics, and more responsiveness toward employment status.

Several possible explanations for the WHOQOL-BREF’s excellent performance include the following. First, low ceiling and floor values can greatly increase the statistical efficiency of detecting differences across health conditions or changes after treatment. The lower ceiling and floor values of the WHOQOL-BREF may partly have resulted from an intrinsically wider range of raw scores for its domains (16 points), compared with the possible ranges of raw scores of the SF-36 domains which are narrower (e.g., they are ≤10 points in Role Physical, Bodily Pain, Social Functioning, and Role Emotional). Second, compared with the SF-36, the scalability and precision of each domain score of

| TABLE 2 Convergent validity: Spearman’s correlation coefficients between the rating scale and domain scores of the WHOQOL-BREF and SF-36 |
|---------------------------------|-----------------|------------------|---------------------|-----------------|-----------------|-----------------|
| Instrument                      | Rating Scale    | Overall Quality of Life and General Health* | Physical Capacity* | Psychological Well-Being* | Social Relationships* | Environment* |
| Rating scale                    | --              | 0.68             | 0.73               | 0.64             | 0.54             | 0.57           |
| SF-36                           | Physical Functioning | 0.71           | 0.57               | 0.78               | 0.57             | 0.50             | 0.54           |
| Role Physical                   | 0.47             | 0.35             | 0.51               | 0.40             | 0.33             | 0.48           | 0.48           |
| Bodily Pain                     | 0.64             | 0.52             | 0.68               | 0.56             | 0.48             | 0.55           |
| General Health                  | 0.72             | 0.65             | 0.69               | 0.62             | 0.45             | 0.59           |
| Vitality                        | 0.59             | 0.59             | 0.67               | 0.65             | 0.48             | 0.62           |
| Social Functioning              | 0.50             | 0.52             | 0.62               | 0.63             | 0.43             | 0.58           |
| Role Emotional                  | 0.32             | 0.30             | 0.41               | 0.37             | 0.24             | 0.39           |
| Mental Health                   | 0.36             | 0.51             | 0.52               | 0.59             | 0.40             | 0.56           |

* The underlined correlation coefficients are assumed to be convergent with a value ≥0.4, and all P values are less than 0.0001.
<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Statistic</th>
<th>WHOQOL-BREF</th>
<th></th>
<th>SF-36</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OQL</td>
<td>PC</td>
<td>PW</td>
<td>SR</td>
<td>EN</td>
<td>PF</td>
<td>RP</td>
<td>BP</td>
<td>GH</td>
<td>VT</td>
<td>SF</td>
<td>RE</td>
<td>MH</td>
</tr>
<tr>
<td>Age (year)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤43</td>
<td>Mean</td>
<td>55.1</td>
<td>59.3</td>
<td>60.9</td>
<td>55.1</td>
<td>66.9</td>
<td>69.7</td>
<td>70.9</td>
<td>55.7</td>
<td>58.7</td>
<td>73.7</td>
<td>74.5</td>
<td>65.1</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>48.9</td>
<td>52.2</td>
<td>51.7</td>
<td>50.2</td>
<td>50.6</td>
<td>54.0</td>
<td>53.8</td>
<td>63.3</td>
<td>48.4</td>
<td>54.8</td>
<td>69.3</td>
<td>68.3</td>
</tr>
<tr>
<td></td>
<td>P value</td>
<td>0.015</td>
<td>0.006</td>
<td>0.094</td>
<td>0.026</td>
<td>0.026</td>
<td>0.015</td>
<td>0.007</td>
<td>0.005</td>
<td>0.008</td>
<td>0.088</td>
<td>0.133</td>
<td>0.259</td>
</tr>
<tr>
<td>Educational level</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Junior high or below</td>
<td>Mean</td>
<td>51.3</td>
<td>55.0</td>
<td>53.2</td>
<td>58.4</td>
<td>52.4</td>
<td>59.0</td>
<td>58.8</td>
<td>67.1</td>
<td>50.5</td>
<td>56.9</td>
<td>71.9</td>
<td>68.9</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>54.2</td>
<td>56.0</td>
<td>54.7</td>
<td>59.7</td>
<td>54.3</td>
<td>65.1</td>
<td>69.8</td>
<td>68.4</td>
<td>56.0</td>
<td>57.0</td>
<td>71.6</td>
<td>77.0</td>
</tr>
<tr>
<td></td>
<td>P value</td>
<td>0.278</td>
<td>0.268</td>
<td>0.498</td>
<td>0.534</td>
<td>0.365</td>
<td>0.269</td>
<td>0.076</td>
<td>0.655</td>
<td>0.050</td>
<td>0.976</td>
<td>0.941</td>
<td>0.631</td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single/widowed/divorced</td>
<td>Mean</td>
<td>48.3</td>
<td>51.1</td>
<td>50.4</td>
<td>53.5</td>
<td>49.9</td>
<td>50.5</td>
<td>57.6</td>
<td>64.0</td>
<td>52.5</td>
<td>53.6</td>
<td>67.4</td>
<td>69.7</td>
</tr>
<tr>
<td>Spouse present</td>
<td>Mean</td>
<td>54.0</td>
<td>58.1</td>
<td>55.1</td>
<td>61.1</td>
<td>54.4</td>
<td>65.6</td>
<td>64.8</td>
<td>69.0</td>
<td>52.5</td>
<td>58.3</td>
<td>73.6</td>
<td>72.6</td>
</tr>
<tr>
<td></td>
<td>P value</td>
<td>0.042</td>
<td>0.013</td>
<td>0.050</td>
<td>0.001</td>
<td>0.041</td>
<td>0.009</td>
<td>0.266</td>
<td>0.101</td>
<td>0.982</td>
<td>0.064</td>
<td>0.058</td>
<td>0.628</td>
</tr>
<tr>
<td>Employment status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>Mean</td>
<td>58.6</td>
<td>64.7</td>
<td>59.9</td>
<td>64.8</td>
<td>57.1</td>
<td>77.9</td>
<td>75.0</td>
<td>74.2</td>
<td>59.2</td>
<td>60.4</td>
<td>79.0</td>
<td>78.2</td>
</tr>
<tr>
<td>No</td>
<td>Mean</td>
<td>47.0</td>
<td>48.6</td>
<td>48.3</td>
<td>53.8</td>
<td>49.5</td>
<td>47.1</td>
<td>52.4</td>
<td>61.8</td>
<td>46.7</td>
<td>53.9</td>
<td>65.5</td>
<td>66.1</td>
</tr>
<tr>
<td></td>
<td>P value</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
</tr>
<tr>
<td>Level of injury</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quadriplegia</td>
<td>Mean</td>
<td>48.5</td>
<td>51.3</td>
<td>50.8</td>
<td>55.2</td>
<td>49.9</td>
<td>56.5</td>
<td>52.4</td>
<td>62.2</td>
<td>50.7</td>
<td>53.4</td>
<td>69.6</td>
<td>62.9</td>
</tr>
<tr>
<td>Paraplegia</td>
<td>Mean</td>
<td>55.8</td>
<td>60.4</td>
<td>56.4</td>
<td>62.2</td>
<td>56.0</td>
<td>65.3</td>
<td>71.8</td>
<td>72.2</td>
<td>54.1</td>
<td>60.1</td>
<td>73.7</td>
<td>79.4</td>
</tr>
<tr>
<td></td>
<td>P value</td>
<td>0.004</td>
<td>0.000</td>
<td>0.011</td>
<td>0.001</td>
<td>0.002</td>
<td>0.096</td>
<td>0.001</td>
<td>0.000</td>
<td>0.199</td>
<td>0.003</td>
<td>0.168</td>
<td>0.002</td>
</tr>
<tr>
<td>Self-care ability</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>Mean</td>
<td>58.7</td>
<td>64.5</td>
<td>58.8</td>
<td>62.8</td>
<td>56.3</td>
<td>79.0</td>
<td>73.5</td>
<td>74.0</td>
<td>58.4</td>
<td>62.6</td>
<td>77.8</td>
<td>81.1</td>
</tr>
<tr>
<td>No</td>
<td>Mean</td>
<td>45.7</td>
<td>47.4</td>
<td>48.4</td>
<td>54.7</td>
<td>47.8</td>
<td>42.3</td>
<td>51.1</td>
<td>60.6</td>
<td>46.5</td>
<td>51.1</td>
<td>65.5</td>
<td>61.9</td>
</tr>
<tr>
<td></td>
<td>P value</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.001</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Abbreviations: PF, Physical Functioning; RP, Role Physical; BP, Bodily Pain; GH, General Health; VT, Vitality; SF, Social Functioning; RE, Role Emotional; MH, Mental Health; OQL, Overall Quality of Life and General Health; PC, Physical Capacity; PW, Psychological Well-Being; SR, Social Relationships; En, Environment. Shaded values indicate differences are significant at $P < 0.05$. 
the WHOQOL-BREF summed from the item scores should have smaller variation because the five response descriptors of each item in the WHOQOL-BREF were selected by Thurstone’s equal-appearing interval scaling procedures. Finally, the domains and items in the WHOQOL-BREF may be more qualitatively sensitive at reflecting themes that represent the HRQL for persons with SCIs. For example, inclusion of the Environmental domain in the WHOQOL-BREF is important for persons with SCIs who have sustained a permanent disability and who need greater physical or social environmental support for their independence. In addition, items in the WHOQOL-BREF’s Social relationships domain contain personal relationships, social support and sexual activity, whereas the SF-36’s Social functioning only emphasizes on social activities.

Despite the conceptually related domains of the WHOQOL-BREF and the SF-36 being moderately or highly convergent, there may be differences in the nature of items of the two measures. In general, items of the WHOQOL-BREF are apparently more subjective compared with the SF-36 items. For example, items in the Physical functioning of the SF-36 tend to assess activity levels of daily living, whereas those in the Physical capacity of the WHOQOL-BREF tend to evaluate the well-being of physical functioning. Using Dijkers’s approach to classifying HRQL instruments into those involving personal achievements, reactions to the achievements, and/or expectations about those achievements, the WHOQOL-BREF and the SF-36 seem to differ in the classification by involving personal expectations and achievements, respectively.

Sensitive questions, such as those on sexual activity in the WHOQOL-BREF, were difficult to ask and respond to by the telephone, and that might have increased the proportion of missing values. High missing rates of items related to sexuality also were found in an elderly population of Taiwan (16.5%), cancer patients in the UK (19%), and patients with chronic liver disease in the Netherlands (12%–21%).; therefore, the nonresponse to such questions might be attributable to the mode of administration, the study population, and/or local cultural characteristics. Although the high missing rate for this item among the subjects does supply some information implying that there is little sexual activity dissatisfaction with current sexual activities or just reluctance to discuss sexual matters, it would be better if that item were to be eliminated or replaced by another item in the Social relationships domain of the full version of the WHOQOL questionnaire in which each facet consists of four items.

There are at least two limitations to this study. First, generalizing the results to all persons with SCIs needs to be done in a mindful way because a substantial number of SCI subjects in the registry were not reached. Despite no statistically significant differences being detected in any of the characteristics recorded in the registry, more of the nonparticipants tended to be tetraplegic and older, indicating a poorer health status, than those participating in the study. In addition, new cases with SCI were not included in the study. The performance of the HRQL measures may be somewhat different altered in population with more-severe or newer SCIs. Second, the validity and reliability of the telephone interviews for the WHOQOL-BREF need to be further examined. The telephone interviews for the SF-36 and other HRQL measures were comparable with those of personal interviews and self-administration. Even though this administration mode might produce socially acceptable responses, it was necessary for us to conduct these interviews by telephone because our subjects...
lived nationwide, and their physical health was such that it was inconvenient for them to leave their homes for personal interviews or to fill in the study questionnaire. In this study, the mail mode may have been inappropriate because a large proportion of caregivers are foreign attendants many of whom are usually unable to read Chinese and would not have been able to assist in marking the answers; furthermore, subjects may have been unwilling to reveal personal perception on some sensitive items to caregivers or family members. The mail mode also was found to have higher missing data rates compared with the telephone interview. 53–55

In this study, the distribution of each of the domain scores, measured by the WHOQOL-BREF and the SF-36, were described, and comprehensive comparisons of validity, reliability, and responsiveness were also made between the two instruments. We found that the psychometric performances of the WHOQOL-BREF at large were comparable with or even better than those of the SF-36 among persons with SCIs. The study suggests that the WHOQOL-BREF is an appropriate generic HRQL instrument for persons who have experienced traumatic SCIs.

REFERENCES

35. Ware JE, Gandek B: Overview of the SF-36 Health Survey and the International Quality of Life Assessment (IQOLA) Project. J Clin Epidemiol 1998;51:903–12
A Prospective Study to Validate an Impairment Questionnaire for Major Trauma Survivors

ABSTRACT

Objective: To develop and validate an impairment questionnaire that will provide an estimate of whole-person impairment in patients who have suffered major trauma.

Design: This was a multicenter prospective study involving a convenience sample of 43 volunteer participants who had sustained major trauma within 1 yr of study commencement. Patients were recruited from two trauma centers in Ontario, Canada. The impairment questionnaire was developed as a self-administered questionnaire based on the American Medical Association’s Guides to the Evaluation of Permanent Impairment, Fourth Edition.

Results: Clinician assessments of whole-person impairment showed adequate interrater (\( r \geq 0.55, P \leq 0.03 \)) and intrarater (\( r \geq 0.62, P \leq 0.055 \)) reliabilities across dimensions. The impairment questionnaire correlated significantly with clinician assessments at the initial visit (\( r \geq 0.57, P < 0.001 \)) and at follow-up (\( r > 0.60, P < 0.001 \)). Comparison of the physical and emotional subcategories of the impairment questionnaire, whole-person impairment ratings by physicians, and the Short Form 36 demonstrated good convergent and divergent validity of the impairment questionnaire.

Conclusions: The impairment questionnaire is a reliable and valid self-administered tool that can be used to evaluate physical impairment after major trauma.

Key Words: Impairment, Multiple Trauma, Disability, Quality of Life
Trauma is a leading cause of mortality and morbidity in those under 40 yrs of age. Recognition of the impact of trauma on society has led to the development of systems of trauma care dedicated to prevention, prehospital, and in-hospital care, as well as the rehabilitation of trauma survivors. These centers have contributed to a decrease in the overall fatality rate attributable to trauma. However, this trend can be expected to lead to an increase in the already large number of trauma survivors who have residual impairments.

Impairment can be evaluated by two general methods, either by simple identification of normality or abnormality of organ function or body part, or by quantification and combination of a number of different deficits. Disability measures, in essence, evaluate whole-person functioning, generally through rating the ability to complete basic self-care activities. Most impairment scales, on the other hand, tend to be disease- or syndrome specific, such as scales for stroke, multiple sclerosis, or coma. With regards to trauma, where many body systems can be affected, the Abbreviated Injury Score and the derived Injury Severity Score (ISS) have often been used to capture severity. It has also been noted that as ISS values increase, longer periods of disability are reported, suggesting that the scale measures function rather than impairment. Efforts have been made to estimate morbidity secondary to trauma with a measure called the Injury Impairment Scale. However, validation and use of this scale has not proceeded to a significant degree since it was introduced in 1989.

To date, long-term studies measuring outcome after trauma have focused on disability scales, and very few have studied impairment as an outcome. The current standard for the assessment of physical impairment is the American Medical Association (AMA) Guides to the Evaluation of Permanent Impairment. The definition used by the AMA for impairment closely parallels that of the World Health Organization. The AMA guides are comprehensive in that 12 bodily systems are evaluated. An impairment percentage may be derived for each system individually or reflected as a percentage of whole-person impairment (WPI). The derived impairment ratings are based only on a medical evaluation of the organ system that is affected.

Although the AMA guides have been considered the gold standard for assessing permanent impairment, there is research need for a cost-effective and time-efficient tool that does not require the time and expertise of a physician to administer. Such an assessment tool would also provide clinicians with a useful screening measure to assess patient-perceived impairments after trauma. Therefore, we developed a self-administered impairment questionnaire (IQ) that would provide an estimate of the amount of physical impairment after major trauma. As a second objective, we assessed the reliability and validity of the IQ to assess WPI in patients who have suffered major trauma.

METHODS
IQ Development

The IQ was modeled on the AMA Guides to the Evaluation of Permanent Impairment, Fourth Edition. The content of the IQ was based on the elements of impairment as organized in the AMA guides. Items were generated to represent each of the content elements: upper extremity, lower extremity, spine and pelvis, respiratory and cardiovascular systems, hematopoietic system, vision, ear–nose–throat, digestive system, urinary and reproductive system, endocrine system, skin, and mental and behavioral disorders. Next, specific algorithms were developed to guide patients to determine level of severity. Specialists’ input into relevant areas was sought to improve the algorithms. Specialists were provided with relevant sections of the AMA guides and were also provided with the proposed algorithm. They were then asked to critique the algorithm and complete a standard questionnaire regarding the guides’ content and structure. As a result, five levels (not impaired, minimally impaired, moderately impaired, severely impaired, and extremely impaired) of severity were derived, accompanied by clear definitions. Also, the IQ was piloted by 10 outpatients undergoing rehabilitation to review their interpretation and responses and to estimate the time required to complete the questionnaire. Final adjustments to the scale were made based on the pilot information and comments received from additional physician reviews. The final version of this self-administered IQ consisted of 13 subcategories and 26 items based on the elements of physical impairment as outlined in the AMA guides (see Appendix published as supplemental material at www.AJPMR.com). It also provided a percentage estimate of WPI similar to the scoring method used in the AMA guides.

Participants

Patients who had sustained major trauma within the year before study commencement were identified through two hospital-based trauma databases. Patients were eligible for the study if the ISS was greater than 12, age was at least 18 yrs, and patients were admitted to hospital for at least 24 hrs. Patients were excluded if they were not able to communicate in English, if they had significant aphasia, were pregnant, had a severe head injury (Glasgow Coma Scale score <9), or were paraplegic or tetraplegic. Eligible patients were then con-
tacted via telephone by the trauma research staff or study physicians. Physicians who participated in the study were either trauma surgeons (physician B, physician C) or a specialist in physical medicine and rehabilitation (physician A).

**Procedure**

Trauma survivors who agreed to participate were evaluated in the outpatient trauma clinics at the participating hospitals. Patients were classified into two groups. Participants who had experienced their trauma more than 6 mos before the initial assessment were seen for a first visit and then again 2 wks later. This group was used to assess test–retest reliability because their findings were not expected to change during the 2-wk time frame. Patients who experienced trauma within 6 mos before the initial assessment were reevaluated 3 mos later. This group was used to evaluate the IQ’s sensitivity to change because their findings were expected to change over the longer time frame. Before the study, the study physicians participated in training sessions to ensure consistency in using the AMA guides. Extensive training course manuals were used in conjunction with clinical scenarios and patient evaluations to standardize patient assessments.

**Outcome Measures**

Data were collected using the following instruments.

**ISS**

The ISS is a measure used to describe patients with multiple injuries. For trauma patients, this scoring system is important for assessing the effectiveness of medical care in reducing morbidity and mortality. The score ranges from 0 to 75, with 75 being the highest attainable score and equated with a nonsurvivable injury. A score of 25 yields minimal mortality risk, and a score of 50 is associated with a 50% mortality risk.8

**FIM**

The FIM is an 18-item scale that assesses functional status and independence in activities of daily living. It is the most widely accepted functional assessment measure in use in the rehabilitation community. The total score, as a measure of independence, ranges from 18 to 126. The FIM has been demonstrated to be reliable and a valid estimate of the degree of disability.15

**Short Form 36**

The Short Form 36 (SF-36) is a multipurpose 36-item scale that assesses quality of life defined by functional health and well-being.16–18 The SF-36 is widely used in surveys of general and specific populations when trying to assess the burden of disease or differentiating health benefits of various treatments.

**Pain Assessment**

Pain was measured using a four-item scale that asked about the worst, least, average, and current pain levels that had been experienced in the past 2 wks.20 Each item was rated on a scale ranging from 0 (no pain) to 10 (pain as intense as the patient could imagine). The composite total of these ratings has been found to have acceptable test–retest reliability in comparison with single-item measures, and comparable sensitivity to change.20

**Initial Assessment**

At the initial assessment, all patients met with the research assistant to complete the necessary questionnaires. The IQ took approximately 5 mins for the participants to complete. Patients also completed a SF-36 and the pain assessment20 administered by the research assistant. Also, at the initial assessment, the clinical exam and determination of WPI using the Guides to the Evaluation of Permanent Impairment, Fourth Edition were completed independently by two study physicians. An evaluation form was used for each participant examination to record the findings. At the initial visit, the FIM21 was completed by a physician.

**Follow-up Assessment**

During the follow-up visit, participants were evaluated by only one of the study physicians to assess WPI. Patients in the first group who had experienced their trauma more than 6 mos before the initial assessment were reevaluated 2 wks after the initial assessment. Patients in the second group who experienced trauma within 6 mos before the initial assessment were reevaluated 3 mos later. The IQ, SF-36, and pain ratings were also readministered at the follow-up visit for both groups of participants.

**Statistical Analyses**

Pearson’s correlation coefficients were used to assess the physicians’ interrater reliabilities of the clinical assessment of WPI. Intraclass correlation coefficients and Bland–Altman plots were used to assess the physician test–retest reliability of the clinical assessment of WPI and the IQ.22,23 Results from participants who experienced their trauma more than 6 mos before the initial assessments were used to evaluate the test–retest reliability of the IQ. Similarly, to evaluate sensitivity to change of the IQ, we used results from participants who had experienced trauma less than
6 mos before. Comparisons of participant’s WPI, SF-36, and pain ratings for those who experienced trauma less than 6 mos before were conducted between time one (initial assessment) and 3 mos later (final assessment) using Pearson’s correlation coefficients. Paired t tests were used to confirm that change occurred in patients who experienced trauma within 6 mos and, conversely, did not occur for patients who experienced trauma more than 6 mos earlier.

Validation of the Questionnaire

Concurrent validity of IQ was determined by comparing the results of the IQ measure with the average of the two physicians’ WPI findings at the initial assessment. Also, the average IQ measure was compared with the WPI measures at the follow-up visit to further assess concurrent validity. Construct validity was determined by Pearson’s correlation coefficient comparing the IQ physical and emotional elements with those of the WPI, SF-36, and the pain assessment. Again, the average of the two physicians’ WPI findings at the initial assessment was used as the value for WPI to determine construct validity. In addition, 13 subcategories of WPI and IQ were compared by means of endorsement rates and tests of the degrees of association, using t tests.

Research ethics board approval was obtained from all participating institutions.

RESULTS

Study Participants

A total of 43 participants (19 females, 24 males) were recruited from the outpatient population of two lead trauma centers in Canada: the Ottawa Hospital General Campus (Ottawa, Ontario) and London Health Sciences Centre (London, Ontario). Twenty-six participants had experienced their trauma more than 6 mos before study onset, and 17 had experienced trauma within 6 mos before the initial assessment. The mean age was 42.5 ± 14.8 (SD) yrs, with a range of 18–78 yrs. Mean ISS was 24.4 ± 11.3 (SD), with a range of 13–75. Mean FIM score at the time of hospital discharge was 104.2 ± 22.5 (SD) and ranged from 53 to 126 (Table 1).

Inter- and Intrarater Reliability of Clinical Assessments

Interrater Reliabilities Pooled by Site

The interrater reliability (physician A and physician B) for Ottawa Trauma more than 6 mos and less than 6 mos before study onset is 0.55, 95% CI (0.06, 0.83), \( P = 0.032 \) (Fig. 1). The interrater reliability (physician A and physician C) for London Trauma more than 6 mos and less than 6 mos before study onset is 0.90, 95% CI (0.79, 0.95), \( P < 0.001 \) (Fig. 1).

Interrater Reliabilities by Site and Trauma

Interrater reliability (physician A and physician B) for Ottawa Trauma more than 6 mos before study onset is 0.69, 95% CI (0.23, 0.90), \( P = 0.001 \) (Fig. 1). However, interrater reliability (physician A and physician B) for Ottawa Trauma less than 6 mos before study onset could not be determined because of an insufficient number of observations (Fig. 1). For London Trauma more than 6 mos before study onset, interrater reliability (physician A and physician C) is 0.90, 95% CI (0.69, 0.97), \( P < 0.001 \) (Fig. 1). Interrater reliability (physician A and physician C) for London Trauma less than 6 mos before study onset is 0.90, 95% CI (0.71, 0.97), \( P < 0.001 \) (Fig. 1).

Intrarater Reliability

The intraclass correlations (ICC) for each rater (physicians A, B, C) at initial and final assessments were calculated. Because individual ICCs seemed

<table>
<thead>
<tr>
<th>TABLE 1 Patient demographics</th>
<th>Trauma &gt;6 mos ((n = 26))</th>
<th>Trauma &lt;6 mos ((n = 17))</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Patient Characteristic</strong> ((n = 43))</td>
<td>((60.5%))</td>
<td>((39.5%))</td>
</tr>
<tr>
<td>Sex (males), (n (%))</td>
<td>24 (55.5%)</td>
<td>11</td>
</tr>
<tr>
<td>Sex (females), (n (%))</td>
<td>19 (44.5%)</td>
<td>15</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>42.5 (14.8)</td>
<td>43.1 (13.6)</td>
</tr>
<tr>
<td>Median (range)</td>
<td>45 (18–78)</td>
<td>(18–78)</td>
</tr>
<tr>
<td>Injury Severity Score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>24.4 (11.3)</td>
<td>23.2 (9.1)</td>
</tr>
<tr>
<td>Median (range)</td>
<td>22 (13–75)</td>
<td>(13–50)</td>
</tr>
<tr>
<td>FIM score at discharge</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>104.2 (22.5)</td>
<td>99.7 (20.5)</td>
</tr>
<tr>
<td>Median (range)</td>
<td>112 (53–126)</td>
<td>(60–126)</td>
</tr>
</tbody>
</table>
consistent, the overall estimate of ICC for rater A and B can be used; 0.62, 95% CI (0.39, 0.79). Bland–Altman plots also demonstrate the test–retest reliability of the clinical assessment at 2 wks for physician A ($P \leq 0.66$) (Fig. 2).

Test–Retest Reliability of IQ

Results from patients who experienced their trauma more than 6 mos before the initial assessment were used to assess test–retest reliability of the IQ (Fig. 3). Comparisons of the WPI (average of two physicians at initial assessment), pain, and SF-36 measures between the initial and final assessments demonstrated no significant changes, with high correlations and nonsignificant differences in mean values. Similar results extracted from these comparisons on the IQ findings supported the test–retest reliability of the IQ ($r = 0.89$, $P < 0.001$ and $t$ test $P$ value $>0.787$) (Table 2).

Test Sensitivity to Change

Results from patients who experienced their trauma within 6 mos before the initial assessment were used to assess the sensitivity of the IQ. The IQ was as sensitive as the WPI, pain assessment, and SF-36 measurements in sensitivity to changes between the initial assessment and the follow-up assessment at 3 mos. Results indicate high correlations ($r \geq 0.71$, $P \leq 0.003$), but also a significant difference of means ($P \leq 0.027$) indicating an improvement in total scores over time (Table 3).

Concurrent Validity of IQ

To assess the concurrent validity of the IQ, results of patients from both trauma groups were pooled so that comparisons could be made with the Physician Whole-Person Impairment ratings, which are considered the gold-standard comparator for this study. Correlations were then computed for each combination of WPI and IQ measurements at the initial and follow-up assessments, and differences in mean values were examined.

Significant correlations were found between the IQ measures and the average of the two physicians' WPI ratings at both the initial assessment ($r \geq 0.57$, $P < 0.0001$) and follow-up assessments ($r > 0.60$, $P < 0.0001$). However, the means of the IQ scores were significantly higher than those of the physicians at both assessment points ($P < 0.001$). According to the same samples, the means of IQ values were higher than the WPI findings. Differences of means were 35.24 at the initial assessment and 37.89 at the follow-up assessment.

Figure 4 shows the relation between the average of two physicians' WPI measures plotted against IQ scores at the initial assessment, and Figure 5 shows WPI measures plotted against IQ...
scores at the follow-up assessment. A positive correlation exists between the IQ scores and WPI measures.

**Construct Validity**

Construct validity provides further support for a measure by demonstrating that it correlates with expected outcomes that would not necessarily be considered a gold standard. Convergent and divergent validity are elements of construct validity because some measures may be expected to correlate highly with an experimental measure (convergent validity), whereas some elements of a measure would be expected to correlate poorly (divergent validity). To explore the construct validity of the IQ, the subcategory scores of the physical and emotional components of the IQ, WPI, and SF-36 were compared. To avoid lack of independence, the data from all patients (both groups combined) at the initial assessment were analyzed, again using the average of two physicians’ measurements as the WPI measurement. Higher correlations were found between the physical elements of the IQ with those of SF-36 and
The measures used in this study attempted to evaluate impairment (IQ, WPI, ISS) and functional independence (FIM) as overall contributors and predictors of health-related quality of life as measured by the SF-36. Correlation of the FIM with the subcategories on the data from the initial assessment were calculated. The means of the IQ and WPI values in the subcategories of upper extremity, lower extremity, respiratory and cardiovascular systems, mental and behavioral disorders were significantly different (P < 0.049). According to the same samples, IQ mean values were considerably higher (upper extremity, 8.7; lower extremity, 11.3; respiratory and cardiovascular systems, 11.1; urinary and reproductive systems, 18; and mental and behavioral disorders, 42.7).

**DISCUSSION**

The IQ seems to have potential as a valid measure to evaluate impairment for major trauma survivors. The IQ was modeled on the AMA’s Guide to the Evaluation of Permanent Impairment, Fourth Edition, and the item content was based on the elements of impairment as organized in the AMA guides. This tool was developed so that it could be administered as a self-report questionnaire and/or used as a screening measure to complement and augment the clinical evaluation of impairment. A previous study showed that the IQ has predictive validity for the evaluation of permanent impairment. The IQ questions of the above-mentioned categories should be modified based on the AMA guides to improve the validity of the IQ.

Further investigation was completed to explore the areas of the IQ that needed to be revised, either in the wording of questions or in the scoring. This analysis was performed on the data retrieved from the initial assessment, using the WPI measures collected by one of the physicians who had examined all 43 patients. Both IQ and WPI data were separated into 13 similar subcategories.

Table 5 summarizes 13 subcategories of WPI and IQ and compares the endorsements of health subcategories by both tools. More than 50% of the cases scored by the physical examinations were endorsed in IQ results, in all categories except for skin. However, the physician endorsed fewer than 50% of the cases recognized as impaired in IQ responses in the categories of lower extremity, spine and pelvis, nervous system, respiratory and cardiovascular systems, hematopoietic system, ENT, and urinary and reproductive systems. Therefore, the IQ questions of the above-mentioned categories should be modified based on the AMA guides to improve the validity of the IQ.

The measures used in this study attempted to evaluate impairment (IQ, WPI, ISS) and functional independence (FIM) as overall contributors and predictors of health-related quality of life as measured by the SF-36. Correlation of the FIM with the subcategories of the WPI and SF-36. Divergent validity is demonstrated by the poorer correlation of physical subcategories compared with emotional subcategories.

The measures used in this study attempted to evaluate impairment (IQ, WPI, ISS) and functional independence (FIM) as overall contributors and predictors of health-related quality of life as measured by the SF-36. Correlation of the FIM with the subcategories of the WPI and SF-36. Divergent validity is demonstrated by the poorer correlation of physical subcategories compared with emotional subcategories.

The measures used in this study attempted to evaluate impairment (IQ, WPI, ISS) and functional independence (FIM) as overall contributors and predictors of health-related quality of life as measured by the SF-36. Correlation of the FIM with the subcategories of the WPI and SF-36. Divergent validity is demonstrated by the poorer correlation of physical subcategories compared with emotional subcategories.

The measures used in this study attempted to evaluate impairment (IQ, WPI, ISS) and functional independence (FIM) as overall contributors and predictors of health-related quality of life as measured by the SF-36. Correlation of the FIM with the subcategories of the WPI and SF-36. Divergent validity is demonstrated by the poorer correlation of physical subcategories compared with emotional subcategories.
focus clinical assessment. In addition to the modeling of the IQ, specific algorithms were developed to guide patients in determining the level of severity associated with specific degrees of impairment.

The AMA’s Guides to the Evaluation of Permanent Impairment, Fourth Edition have historically been the gold standard for measuring impairment. These guides have been applied as a reliable tool in clinical research studies such as functional outcome after tibial fractures, assessment and management of upper- and lower-extremity impairment and disability, effects of physiotherapy and occupational therapy in reflex sympathetic dystrophy, the benefit of bilateral stapedotomy, visual function among patients with glaucoma, and assessment of hand impairment.

A number of studies have tested the validity and reliability of the AMA guides’ impairment-evaluation methods. From these studies, two main observations have been identified. First, the AMA guides’ internal deficiencies and the way in which workers’ compensation systems use the ratings for insurance awards suggest that the AMA guides are best used for evaluation of permanent impairment rather than disability. In fact, the AMA guides clearly state that impairment ratings should not be directly used to estimate disability but that these ratings are, in fact, the first step in determining disability. Other factors such as age, education, premorbid functioning, family and community support, and personal and community resources must also be determinants when applying impair-

<table>
<thead>
<tr>
<th></th>
<th>IQP</th>
<th>WPI P</th>
<th>SF-36P</th>
<th>IQE</th>
<th>WPI E</th>
</tr>
</thead>
<tbody>
<tr>
<td>IQP</td>
<td>0.67 (0.46, 0.81)*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WPI P</td>
<td>0.62 (−0.77, −0.39)*</td>
<td>−0.35 (−0.59, −0.05)*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF-36P</td>
<td></td>
<td>−0.07 (−0.37, 0.23)</td>
<td></td>
<td>−0.37 (−0.60, −0.08)*</td>
<td></td>
</tr>
<tr>
<td>IQE</td>
<td>0.20 (−0.11, 0.47)</td>
<td></td>
<td>−0.02 (−0.32, 0.28)</td>
<td></td>
<td>0.57 (0.33, 0.75)*</td>
</tr>
<tr>
<td>WPI E</td>
<td>0.24 (−0.06, 0.51)</td>
<td></td>
<td>−0.25 (−0.51, 0.05)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF-36E</td>
<td>−0.11 (−0.40, 0.20)</td>
<td>0.09 (−0.22, 0.38)</td>
<td>0.24 (−0.06, 0.51)</td>
<td>−0.57 (−0.74, −0.32)*</td>
<td>−0.65 (−0.80, −0.44)*</td>
</tr>
</tbody>
</table>

* Statistically significant at \( \alpha = 0.05 \).
ment ratings to disability evaluations.\(^9,33\) Second, there is a lack of clarity defining which evaluation system to use, either anatomical or range of motion, in presenting situations. For example, the range of motion–based impairment model was introduced as a technically complicated method\(^34\) and was not suitable as the sole determinant of low-back pathology diagnosis.\(^35\) Particularly when aiming to measure or compensate disability.\(^36\) However, to measure impairments after lower-extremity fractures, use of the AMA guides’ anatomical approach has been suggested.\(^32\) Furthermore, in a study measuring impairment us-

![Figure 5](image_url)

**FIGURE 5** Final IQ vs. average of two physicians WPI measures for trauma >6 mos and trauma <6 mos before study onset.

<table>
<thead>
<tr>
<th>Body Part or System</th>
<th>n</th>
<th>Consensus by IQ and WPI</th>
<th>WPI Endorsed (Physician Rating)</th>
<th>IQ Endorsed (Patient Rating)</th>
<th>Con/WPI</th>
<th>Con/IQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper extremity</td>
<td>43</td>
<td>18</td>
<td>18</td>
<td>32</td>
<td>1</td>
<td>0.56</td>
</tr>
<tr>
<td>Lower extremity</td>
<td>43</td>
<td>10</td>
<td>11</td>
<td>25</td>
<td>0.91</td>
<td>0.40</td>
</tr>
<tr>
<td>Spine and pelvis</td>
<td>43</td>
<td>7</td>
<td>8</td>
<td>21</td>
<td>0.88</td>
<td>0.33</td>
</tr>
<tr>
<td>Nervous system</td>
<td>43</td>
<td>3</td>
<td>6</td>
<td>13</td>
<td>0.50</td>
<td>0.23</td>
</tr>
<tr>
<td>Respiratory and cardiovascular system</td>
<td>43</td>
<td>3</td>
<td>3</td>
<td>9</td>
<td>1</td>
<td>0.33</td>
</tr>
<tr>
<td>Hematopoietic system</td>
<td>43</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>NA</td>
<td>0</td>
</tr>
<tr>
<td>Vision</td>
<td>43</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Nose and throat</td>
<td>43</td>
<td>6</td>
<td>10</td>
<td>18</td>
<td>0.60</td>
<td>0.33</td>
</tr>
<tr>
<td>Digestive system</td>
<td>43</td>
<td>6</td>
<td>6</td>
<td>7</td>
<td>1</td>
<td>0.86</td>
</tr>
<tr>
<td>Urinary and reproductive system</td>
<td>43</td>
<td>6</td>
<td>6</td>
<td>14</td>
<td>1</td>
<td>0.43</td>
</tr>
<tr>
<td>Endocrine system</td>
<td>43</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>0.50</td>
<td>0.67</td>
</tr>
<tr>
<td>Skin</td>
<td>43</td>
<td>2</td>
<td>11</td>
<td>4</td>
<td>0.18</td>
<td>0.50</td>
</tr>
<tr>
<td>Mental and behavioral</td>
<td>43</td>
<td>14</td>
<td>16</td>
<td>20</td>
<td>0.88</td>
<td>0.70</td>
</tr>
</tbody>
</table>

*IQ, impairment questionnaire; WPI, whole-person impairment; Con, consensus ratio.*
ing the AMA guides in patients with a fracture of the lower extremity, the anatomical approach provided a more sensitive measure than did the range-of-motion approach.32 These observations were applied in both modeling of the IQ and in clinical measurements of the present study.

Impairment describes physical, cognitive, and emotional abnormalities. After major trauma, multiple organ systems can be involved; however, with recovery, impairment and function will gradually improve. Although impairments may persist, this does not necessarily manifest as functional losses or disability for routine self-care tasks that are measured by the FIM instrument. In fact, for this study, the final FIM score for most patients was near the maximum score of 126 (average 123.8 ± 4.8). This finding is not unreasonable, because all patients were at the outpatient level, and the FIM is generally used as an inpatient rehabilitation outcome measure. However, health-related quality of life as measured by the SF-36 was still appreciably affected for most patients and had a high correlation with the IQ score. Although the primary purpose of this study was to develop and validate an impairment measure, the results suggest that impairment up to 1 yr after trauma remains a factor affecting health-related quality of life for major trauma survivors. These findings coincide with a study that investigated the prevalence of impairments and disabilities after severe multiple trauma.11 At 3-yr follow-up, 80% of patients were still experiencing residual impairments.11 Approximately 74% experienced physical impairments, cognitive impairments were attributed to 32% of the participants, and 25% experienced a decrease in the quality of their social network and an increase in their feelings of loneliness after multiple trauma.11

Because trauma patients can experience longer-lasting residual impairments, the IQ could be used to capture measures of impairment in longer-term studies because of its cost-effective and time-efficient advantages. Another advantage of the IQ as a self-administering tool is its ease of use, even for patients with significant impairment, as this study has shown. The IQ is also reliable when tested over time and between observers, and it is valid when compared with specific domains of other instruments and physician evaluations. Finally, validation of this tool enables researchers and clinicians to estimate physical impairment in an efficient, reliable manner.

We found that when comparing the concordance between two independent physicians estimating WPI, the questionnaire provides a similar level of correlation, which suggests good concurrent validity of the measure. Although the ratings for the IQ categories as rated by the patient were consistently higher than the WPI ratings of the physician, especially for certain dimensions, the high correlations suggest that each is rating impairment. Despite the discrepancies between the actual impairment measures, it is important to recognize the value of the good concurrent validity of the IQ and its potential usefulness, both for screening purposes in clinician ratings of impairment and for research. It may be necessary to revise the scaling of the IQ in future editions to maximize the concordance with physician ratings in terms of absolute percentage impairment. Construct validity of the IQ has been demonstrated in that it has reasonable correlation with other measures such as the SF-36, which has overlap of measured items. These results suggest that it can be used as an independent measure of impairment.

Although the IQ does demonstrate reliability and validity as an impairment measure for persons who have had major trauma, there were limitations to this study. The questionnaire itself is relatively long and may be more difficult to use for persons with lower educational levels. Some of the subcategories used to estimate WPI by both physicians and the IQ were less well represented, including respiratory and cardiovascular systems, hematopoietic system, vision, and endocrine system. In essence, these subcategories may not contribute to the overall score, and their reliability for the IQ might be questionable. A larger population should be studied, or the possibility of eliminating these items in assessing the major trauma population should be considered. Because the overall age tends to be younger for the major trauma population, exclusion of some subcategories may be necessary because impairments may be encountered infrequently. The smaller sample size used in this study might indicate that not all trauma types have been captured and that the precision of estimates could be less than accurate.

Future study should explore the relationship of impairment to health-related quality of life and outcome of major trauma survivors in a prospective fashion. The IQ may be used as a self-administered questionnaire that can estimate WPI and facilitate the study of larger populations of trauma survivors. Rasch analysis37 and other probabilistic models38 can possibly facilitate a more informed evaluation and development of the IQ. This remains a goal for future research, along with studies involving a large number of both patients and raters to better enable such modeling.

The validity of the IQ for other populations will also need to be studied.

CONCLUSIONS

The IQ seems to have potential as a self-administered outcome measure to evaluate patient-perceived impairments after major trauma. This
study demonstrates that the IQ seems to validly measure impairment. Further prospective study will be required to confirm the reliability of the IQ and its potential generalizability for use in other settings.

ACKNOWLEDGMENTS

Sincere gratitude is expressed to Dorothyann Curran and Lynn Macleay for technical assistance in preparation of the IQ.

REFERENCES

2. Trunkey DD: Trauma. Accidental and intentional injuries account for more years of life lost in the U.S. than cancer and heart disease. Among the prescribed remedies are improved preventive efforts, speedier surgery and further research. Sci Am 1983;249:28–35
14. Deleted in proof
19. Deleted in proof
Falls and Gait Characteristics Among Older Persons with Peripheral Neuropathy

ABSTRACT

Objective: To prospectively determine the frequency and circumstances of falls in older persons with peripheral neuropathy and to identify gait characteristics on smooth and irregular surfaces associated with falls in this same population.

Design: This was a descriptive and observational study of a prospective group cohort. Spatial and temporal gait measures on smooth and irregular surfaces, as well as basic demographic and clinical data, were obtained in 20 older persons with peripheral neuropathy. Falls and fall-related injuries were then prospectively determined for 1 yr.

Results: Thirteen of 20 (65%) subjects fell, and 6 of 20 (30%) subjects sustained a fall-related injury during the year of observation. Of the 76 reported falls, 69 (90.8%) were associated with a surface abnormality (irregular or slick). Gait measures on the smooth surface did not distinguish between fall groups. On the irregular surface, however, step-time variability tended to be higher for those subjects who fell than for those who did not (89 ± 29 vs. 64 ± 26 msecs, respectively; \( P = 0.077 \)) and for those who were injured from a fall compared with those who were not injured (101 ± 21 vs. 71 ± 29 msecs, respectively; \( P = 0.038 \)).

Conclusions: Older patients with peripheral neuropathy have a high rate of falls, and these falls are often associated with walking on irregular surfaces. Gait analysis on an irregular surface may be superior to that on a smooth surface for detecting fall risk in this patient population.

Key Words: Gait, Neuropathy, Environment, Accidental Falls, Injury
It has been well documented that falls pose a significant threat to the health and quality of life of older persons, with approximately one third of those living in the community falling annually. Injuries resulting from falls cost $6 billion per year in the United States. Therefore, efforts have been made to identify the characteristics placing an individual at increased risk for falling.

Previous work has found that older persons with a distal, symmetric, sensorimotor peripheral neuropathy (PN), which causes a distal to proximal loss of sensation and strength, have increased (less precise) ankle proprioceptive thresholds and decreased lower-extremity strength and rates of strength development compared with similarly aged controls without PN. These sensory and motor impairments can predispose patients with PN to mediolateral instability and falls. Accordingly, subjects with PN show increased frontal-plane center-of-pressure displacement during stance, and several studies have indicated that the presence of PN increases fall risk among older and younger persons and places them at greater risk of injury from a fall. However, studies that have analyzed groups of subjects selected for the presence of PN were retrospective and, therefore, may have underestimated the incidence of falls related to PN. In support of this, Cummings et al. showed that recall of falls among older persons was associated with errors in the range of 13–33%. Therefore, the first goal of this research was to perform a descriptive, prospective study of older subjects with PN to more accurately characterize the frequency and circumstances of falls in this high-risk population.

Although it is clear that PN increases fall risk, not all persons with PN do fall. Previous work demonstrated that more severe PN and increased body mass index (BMI) were associated with increased fall risk among older persons with PN. However, the ability of gait analysis to predict falls within this high-risk group has not, to our knowledge, been explored. It has been our clinical experience that older persons with PN have particular difficulty with the combination of an irregular walking surface and low light conditions. This finding is in line with the known distal impairments in ankle sensory and motor function and with an increased reliance on vision in the absence of reliable somatosensory information. In previous work, we have found that gait analysis on an irregular surface has been superior to similar analysis on a smooth surface at differentiating healthy young from healthy old and healthy old from older subjects with PN. Similarly, Menz et al. found that an irregular surface accentuated gait differences between subjects with and without PN and with and without physiologic factors known to be associated with fall risk. Therefore, the second goal of our study was to determine the extent to which gait measures on smooth and irregular surfaces were prospectively associated with falls and injurious falls. The primary hypothesis was that a measure of gait variability on the irregular surface, either of step width or step time, would be prospectively associated with falls.

**METHODS**

**Subjects**

Subjects were recruited from the University of Michigan Electrodiagnostic Laboratory and the Physical Medicine and Rehabilitation Outpatient Orthotics and Prosthetics Clinic and participated in a previous study investigating the effect of interventions on gait variability. All patients underwent history, physical examination, and electrodiagnostic testing. The project was given approval by the University of Michigan institutional review board, and all subjects gave written informed consent.

Inclusion criteria were age between 50 and 80 yrs, ability to speak and understand English, and ability to ambulate household distances without an assistive device. Subjects also met criteria for a distal, symmetric sensorimotor PN by the presence of (1) symmetric symptoms consistent with PN, (2) a physical examination consistent with PN (symmetrically absent or relatively decreased Achilles reflexes, decreased distal lower-extremity sensation that improved proximally and normalized proximal to the knees), and (3) electrodiagnostic evidence consistent with a distal, symmetric, sensorimotor polyneuropathy in that one or more abnormalities were seen in the peroneal motor and sural responses. All subjects demonstrated sural responses that were absent or of decreased amplitude (<6 µV) and peroneal motor responses that were of decreased amplitude (<2.0 mV) and/or conduction velocity (<41.0 m/sec).

Exclusion criteria were subject report of abnormal vision despite correction, subject report of change in health status (e.g., joint replacement or vascular event) between the time of evaluation and falls data recording, weight greater than 136 kg (300 pounds), evidence on physical examination of central neurologic dysfunction, or musculoskeletal abnormality such as severe scoliosis or amputation.

**Clinical Evaluation**

The clinical evaluation included the recording of height, weight, and medications used, with special note of medications known to be associated with falls, such as benzodiazepines, tricyclic antidepressants, and antihypertensives.

Additional
evaluation included muscle stretch reflexes, lower-extremity sensation (vibration, proprioception, monofilament testing, pinprick), and strength of the great-toe extensors, ankle dorsiflexors, and hand dorsal interossei (manual muscle testing). The reflex, sensation and strength tests were used to determine the Michigan Diabetes Neuropathy Score (MDNS). The MDNS was used as a clinical measure of PN severity and is a 0- to 46-point scale (higher score reflecting more severe PN) that correlates well with more extensive neuropathy staging scales. In addition, each subject underwent a test of unipedal stance time with the mean of three trials used as the subject’s unipedal stance time. Prospective data collection began 11.5 ± 4.5 mos (range 5–19 mos) after the subjects had undergone electodiagnostic and clinical evaluation.

Prospective Falls Data Collection

Each subject was given 26 bimonthly calendars, each of which consisted of a daily log for a 2-wk span. Each day, the subject checked a box to indicate whether he or she had experienced any falls. Space was also available for subjects to comment on the nature of the falls and under what circumstances they had occurred. At the end of each 2-wk period, subjects returned the surveys in a preaddressed and stamped envelope. If no response was received, or if a fall was indicated, the subject was contacted by the research coordinator (TD) for further details. By combining a daily log into a calendar to be returned twice a month, it was hoped that recall bias would be minimized. The as-needed follow-up interviews, as suggested by Tinetti et al., were planned to improve details of any falls and to increase compliance with the fall diaries while minimizing intrusion on the subjects. A fall was defined as an unintentional change in body posture that resulted in the subject coming to rest on the ground or other lower level that was not a consequence of a physical blow or loss of consciousness.

Subject Preparation and Experimental Apparatus

These methods have been used in previous work and are described in greater detail elsewhere. The subjects wore flat-soled athletic shoes supplied by the laboratory. Subjects were allowed 5 mins to accommodate to the new shoes by walking at their leisure in a flat, well-lit hallway adjacent to the laboratory. The subjects were placed in a safety harness that was secured to an overhead track. The harness suspension was adjusted to prevent the knees from coming into contact with the floor when the subject hung unsupported. For all trials, the subjects were instructed to walk at their own pace, as if they were walking to mail a letter. The subjects performed 10 trials (two lengths of the walking surface = one trial) on the smooth surface and then 10 trials on the irregular surface. Subjects were given 2 mins of rest after the first five trials in each environment and 5 mins of rest after the initial 10 trials.

To create an irregular surface, a 1.5- by 10-m piece of industrial carpet was modified by randomly arranging prism-shaped pieces of wood (height = 1.5 cm, width = 3.5 cm, length = 6–16 cm) beneath the middle 6.5-m section of the carpet at a density of 26 pieces per meter² (Fig. 1). Low light conditions (average 50 lux) were maintained by dimming lights in the room. Two optoelectronic markers (infrared-emitting diodes) were placed 5 cm apart on a malleable aluminum strip (10 × 1.5 cm) inserted under the tongue of each shoe. The top marker was located anterior to the center of the malleoli. A marker was also placed on a belt in the midline at the level of the umbilicus. Two foot switches, each a force-sensing resistor, were placed underneath the insole of each shoe. One switch was placed under the first metatarsophalangeal joint, and the other was placed beneath the calcaneus. Double support was defined as the period of time in the gait cycle during which at least one switch inside each shoe was activated. Kinematic data were measured at 100 Hz using an optoelectronic camera system (OPTOTrack; Northern Digital Corporation, Waterloo, Canada) toward which the subject walked within the boundaries of the walkway.

Gait Data and Statistical Analysis

The kinematic and force data were processed using a custom algorithm to quantify step width, step length, and walking speed. Speed was calculated by taking the time derivative of the waist marker during what was defined as the comfortable gait-speed interval. This interval was found by excluding data taken when the waist velocity was less than 85% of the maximum velocity for that trial. This was done to eliminate steps taken while the subject accelerated to and decelerated from the comfortable gait speed. Similarly, the other gait parameters were only included in the analysis during this interval. Step time was determined by calculating the time elapsed between closure of the right and left metatarsal foot switch during comfortable gait speed. Step width and step length were defined as shown in Figure 1.

SPSS version 12.0 (SPSS Inc., Chicago, IL) was used for all analyses. Descriptive statistics were generated for clinical and demographic data, and gait parameters were determined for subjects on both surfaces. The standard deviations of step width and step time were used as measures of...
step-width and step-time variability. The primary hypothesis that there would be fall-group differences in gait parameters (specifically, variability) was investigated with standard two-tailed t tests. A P value of <0.05 was considered significant, and a P value between 0.05 and 0.10 was considered a trend. The clinical and demographic data were also explored, secondarily, to ascertain that there were no fall-group differences in these areas.

RESULTS
Characteristics of Subjects and Circumstances of Falls

Forty-six subjects with PN were approached for participation. Of these, 12 declined to participate because of self-reported poor health; seven declined to participate because of time constraints; four initially agreed to participate but failed to return the surveys; and three did not respond to the initial request. The remaining 20 subjects (12 women) were included in the analysis. Data were included for the entire 12-mo period for 19 subjects and were included for only 6 mos for one subject who sustained a serious fall injury. The mean ± SD for age, MDNS, height, weight, and BMI were 67.1 ± 9.6 yrs, 15.9 ± 5.3 points, 169.9 ± 11.7 cm, 90.9 ± 25.2 kg, and 31.4 ± 8.0 kg/m², respectively. Causes of PN were diabetes mellitus (11 subjects), connective-tissue disease (2 subjects), familial (1 subject), vitamin B12 deficiency and/or hypothyroidism (1 subject), autoimmune disease related to chronic hepatitis C (1 subject), and idiopathic (4 subjects).

A total of 76 falls were reported. Thirteen subjects (65%) fell at least once, and six subjects (30%) were injured from a fall during the 1-yr period of study. Of the 76 reported falls, 60 (78.9%) occurred while walking on an abnormal surface: nine (11.8%) were described as slips on
an icy or wet surface, and 51 (67.1%) were attributed to a surface irregularity. Sixteen falls (22.5%) occurred while turning or reaching beyond the base of support. Of these, nine (11.8%) occurred in conjunction with a surface irregularity. Therefore, 69 of the 76 reported falls (90.8%) were associated with a surface abnormality, and the majority of these were surface irregularities. None of the subjects reported the use of an assistive device at the time of a fall.

**Clinical Characteristics by Fall Group**

Comparison of subjects who had fallen vs. those who had not fallen revealed no significant differences in MDNS, age, BMI, or unipedal stance time (Table 1). Contrary to the usual finding, the differences in MDNS, age, BMI, or unipedal stance time (Table 1). Contrary to the usual finding, the differences in MDNS, age, BMI, or unipedal stance time (Table 1). Contrary to the usual finding, the differences in MDNS, age, BMI, or unipedal stance time (Table 1). Contrary to the usual finding, the differences in MDNS, age, BMI, or unipedal stance time (Table 1). Contrary to the usual finding, the differences in MDNS, age, BMI, or unipedal stance time (Table 1). Contrary to the usual finding, the differences in MDNS, age, BMI, or unipedal stance time (Table 1). Contrary to the usual finding, the differences in MDNS, age, BMI, or unipedal stance time (Table 1). Contrary to the usual finding, the differences in MDNS, age, BMI, or unipedal stance time (Table 1).

<table>
<thead>
<tr>
<th>Table 1 Demographic variables by fall-group category</th>
<th>Fall (−)</th>
<th>Fall (+)</th>
<th>Injury (−)</th>
<th>Injury (+)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, yrs</td>
<td>67.1 ± 9.6</td>
<td>66.9 ± 10.6</td>
<td>67.5 ± 9.5</td>
<td>65.8 ± 10.8</td>
</tr>
<tr>
<td>Body mass index</td>
<td>32.6 ± 9.9</td>
<td>30.3 ± 7.1</td>
<td>31.6 ± 8.2</td>
<td>30.0 ± 8.0</td>
</tr>
<tr>
<td>MDNS</td>
<td>15.5 ± 7.1</td>
<td>16.2 ± 4.4</td>
<td>16.8 ± 6.1</td>
<td>13.8 ± 1.5</td>
</tr>
<tr>
<td>Unipedal stance time, secs</td>
<td>5.2 ± 6.6</td>
<td>5.8 ± 12.3</td>
<td>4.3 ± 5.6</td>
<td>8.6 ± 17.9</td>
</tr>
<tr>
<td>Medications</td>
<td>3.6 ± 1.9</td>
<td>1.8 ± 1.5</td>
<td>2.8 ± 1.9</td>
<td>1.7 ± 1.4</td>
</tr>
<tr>
<td>Gender, % female</td>
<td>57.1</td>
<td>61.5</td>
<td>50.0</td>
<td>83.3</td>
</tr>
</tbody>
</table>

| Table 2 Spatial and temporal gait measures by fall-group category on smooth and irregular surfaces |
|-----------------------------------------------------|---------|---------|-----------|-----------|
| Step width, mm*                                      | Fall (−) | Fall (+) | Injury (−) | Injury (+) |
| Smooth                                              | 195 ± 41 | 181 ± 30 | 188 ± 35 | 181 ± 31 |
| Irregular                                            | 204 ± 60 | 202 ± 38 | 201 ± 50 | 206 ± 34 |
| Step length, mm                                      | Fall (−) | Fall (+) | Injury (−) | Injury (+) |
| Smooth                                              | 492 ± 102 | 504 ± 81 | 501 ± 100 | 498 ± 49 |
| Irregular                                            | 463 ± 104 | 459 ± 114 | 470 ± 126 | 440 ± 48 |
| Speed, m/sec                                         | Fall (−) | Fall (+) | Injury (−) | Injury (+) |
| Smooth                                              | 0.89 ± 0.22 | 0.87 ± 0.16 | 0.89 ± 0.20 | 0.86 ± 0.12 |
| Irregular                                            | 0.81 ± 0.24 | 0.76 ± 0.25 | 0.82 ± 0.28 | 0.69 ± 0.11 |

**Gait Characteristics of Subjects by Fall Group**

There were no fall-group or fall injury–group differences in any of the gait measures on the smooth surface. However, on the irregular surface there was a trend toward the group that fell demonstrating increased step-time variability. When subjects who sustained a fall-related injury were compared with the remainder of the subjects, the injured subjects demonstrated significantly increased step-time variability on the irregular surface (Table 2). Given that a perfectly metronomic gait would have a step-time variability of zero, the data suggest that the rhythm of the steps of the subjects who fell or who were injured by a fall was more severely disrupted by the irregular surface than that of the subjects who did not fall or who were not injured. Two data-point outliers were identified, one for step-time variability on the smooth surface and the other for step-width variability on the irregular surface. Significance was not affected in either case (Table 2).
DISCUSSION

Frequency of Falls and Injurious Falls

The percentage of subjects reporting a fall during the year of observation (65%) is considerably higher than fall rates typically reported in studies of the general older population. This higher percentage of falls is consistent with previous work that revealed annual fall percentages retrospectively determined, of 55% and 56% in older persons with PN.5,9,10 Taken together, the data from these studies suggest that the annual rate of falls for older persons with PN is approximately double the often-quoted figure of 30% for older persons living in the community.2 This difference is likely attributable to the fact that all of the PN subjects studied are burdened by neuropathy-associated impairments in ankle proprioception4,5 and rate of torque development,6,7 with resultant reduction in balance.29,30 It is assumed that the majority of the comparison populations in the previously referenced studies do not have similarly severe impairments and, thus, fall less frequently. Furthermore, it is the authors’ experience that older persons with neuropathy vary in their perceptions of their limitations, possibly because of the insidious onset of the impairments, and continue to remain active in the community, often exposing themselves to high-risk situations. This exposure is sometimes physician recommended because of the role of walking as exercise to assist in controlling metabolic derangements associated with diabetes mellitus. Similarly, the percentage of subjects injured by falls in this study (30%) is greater than that reported for unselected older populations.31 Six of the 76 falls (approximately 8%) reported in this study led to injury; this ratio of injury to falls is similar to the one in ten reported in larger studies,32 suggesting that older persons with PN are neither more nor less likely to be injured by any single fall than the general population. Therefore, older persons with PN seem to be more frequently injured by falls simply because of their increased frequency of falling.

Influence of Surface Abnormalities on Falls

The overwhelming majority of reported falls (>90%) were associated with an abnormal surface, and the majority of the surface abnormalities were irregularities in surface contour, as opposed to a slick or wet surface. The percentage of falls associated with an irregular surface (78.9%) among the PN population studied seems to be significantly greater than the percentage of falls associated with an irregular surface in an unselected older population (24%).33 The previously identified PN-associated afferent and efferent impairments at the ankle4–7 likely rendered the population studied particularly susceptible to surface irregularities and may explain this difference. Regardless, the data from the present study clearly suggest that older persons with PN must increase their vigilance and exercise caution when on an irregular surface. Touch of a vertical surface or the use of a cane or orthoses that stabilize the ankles in the frontal plane may be useful interventions because they improved spatial and temporal gait regularity in older PN subjects when walking on an irregular surface under low-light conditions.21

Clinical Characteristics

There were no differences between subjects who fell or who were injured by falls vs. those who did not with regard to age, BMI, one-legged balance time, or gender. The absence of a relationship between falls and neuropathy severity or one-legged balance time is counterintuitive, but it may be explained by the small sample size. Alternatively, this may be explained by the fact that the majority of the subjects who fell or who were injured by falls were women. In previous work involving a greater number of older persons with PN, women, unlike men, were not found to have increasing severity of PN or decreased one-legged balance time as fall-risk factors.15

Gait Analysis

PN subjects who fell and/or were injured by falls demonstrated no differences in gait measures on the smooth surface. However, on the irregular surface, subjects who fell demonstrated a trend toward increased step-time variability compared with those who did not fall, and subjects who experienced fall-related injuries demonstrated a significant increase in step-time variability compared with those who were not injured. This finding is of interest for three reasons: (1) the laboratory condition that differentiated between the fall groups mimicked the surface abnormalities subjects often described in association with their falls; (2) the data add to other work that has found that an irregular surface amplifies or identifies gait differences between distinct clinical populations with different clinical backgrounds17–20; and (3) the data coincide with other work suggesting that gait variability in general and step-time variability in particular are associated with increased fall risk among older populations.34,35

Limitations

The most prominent limitation of this study is the sample size. The small number of subjects likely does not adequately represent all older persons with PN and, moreover, precludes the use of a more powerful analytic technique such as logistic
regression that could identify independent predictors of fall risk. In addition, the method of determining the presence of falls may have positively influenced fall reporting in that subjects were rewarded for reporting a fall by receiving a telephone call from an interested person, the study coordinator. The time that elapsed between the electrodiagnostic and clinical evaluation of neuropathy and the collection of data was sufficient that unaccounted-for progression of neuropathy was possible. Finally, although the authors have used the MDNS in subjects with a variety of causes of distal, symmetric, and sensorimotor polyneuropathy and have found excellent correlations with gait measures on smooth and irregular surfaces, the MDNS has not been validated in patients with distal sensorimotor polyneuropathy attributable to causes other than diabetes mellitus.

CONCLUSIONS

Although the study’s sample size is a limitation that dilutes the strength of potential conclusions, the concordance of this study’s data with other work allows some comment. The prospectively determined annual fall rate of the subjects reported in this study is similar to fall rates found in previous retrospective studies of older persons with PN, suggesting that more than half of such patients fall annually. This rate is markedly greater than the 30% annual fall rate reported for older persons in general. The finding that the great majority of falls were associated with irregular surfaces resonates with our clinical experience and previous biomechanical research, which identifies distal impairments in lower-extremity sensory and motor function that would be expected to reduce stability on irregular surfaces. In aggregate, these data suggest that older persons with PN exercise caution and vigilance on irregular surfaces. The increased temporal gait variability on the irregular surface among those with falls and fall-related injuries adds to other work suggesting that gait analysis on an irregular surface may be superior to that on a smooth surface for detecting fall risk among older persons with PN. This finding deserves further exploration.

ACKNOWLEDGMENTS

This study was supported by Public Health Service grants K23AG00989 and P60AG08808

REFERENCES

25. Hurvitz EA, Richardson JK, Werner RA, Ruhl A, Dixon M:


The Effects of Short-Term Exercise Intervention on Falls Self-Efficacy and the Relationship between Changes in Physical Function and Falls Self-Efficacy in Japanese Older People

A Randomized Controlled Trial

ABSTRACT


Objective: To evaluate the effects of short-term exercise intervention on falls self-efficacy and to evaluate the relationships between baseline falls self-efficacy and changes in physical function in older people.

Design: Single-blinded randomized controlled trial. The participants were 171 subjects aged 65 and older. They were randomly assigned into an exercise intervention group or a health education group. The subjects in the exercise intervention group performed an exercise program for 3 mos. Falls self-efficacy was measured using the falls efficacy scale (FES). The measurements of physical function included static and dynamic balance, walking velocity, flexibility, and strength.

Results: There was no significant improvement of FES in either group. But there were significant negative correlations between baseline FES score and the change in maximum walking velocity ($r = -0.29$, $P < 0.018$) and knee extensor strength ($r = -0.25$, $P < 0.040$). Linear regression analysis showed that the change in static balance was related to baseline FES.

Conclusions: The results suggest that a short-term exercise intervention had no effect, possibly because of the high baseline FES scores of the participants, on the confidence of community-dwelling older persons. However, the negative association between FES score and increases in some measures of function suggest that short-term exercise may be beneficial to a subset of older persons with lower FES scores.

Key Words: Falls Self-Efficacy, Physical Function, Short-Term Exercise Intervention, Community-Dwelling Older People
Preventing falls among older people is a major public health issue in many countries. Many researchers have pointed out the relationship between falls and the decline in physical fitness, such as in the balance function, in older people. The fear of falling and a lack of falls self-efficacy, which is a person’s confidence that he or she can perform various activities without falling, are thought to be risk factors for falls as well as decrements of physical capacity. Cross-sectional and prospective studies have revealed that the fear of falling is associated with declines in physical performance, perceived physical function, and activities of daily living. Fear of falling can have negative effects on measures of physical performance, because it can cause apprehension and anxiety that interfere with an individual’s capacity to maximally challenge his or her abilities. Fear of falling commonly arises after a fall has occurred, but such fear can also be present without such a history. If the relationship between physical factors and psychological factors were reciprocal, exercise intervention might provide physical and psychological benefits for older people. If an intervention can improve an individual’s self-efficacy during the intervention period, the physical function of that person may be improved, and it might be possible to increase their activity level.

A number of research studies aimed at reducing the risk of falling in older people have been performed. Most of these studies have examined the effects of exercise intervention programs that lasted 6 mos or longer. When considering exercise intervention as a preventive strategy, it is important to remember that there are many older people in a community. We could not provide sufficient attention to all of our subjects for longer period of time because there are limitations to the services that can be provided in a community, such as the use of specialized staff, and there also are limitations to the social security payments that can cover the cost of such staff.

Some researchers pointed out the role of self-efficacy as a mediator between the fear of falling and improvements in physical function. It should be determined how to maximize the effectiveness of short-term exercise intervention. Therefore, we should investigate whether falls self-efficacy of the subjects might be a result of short-term exercise intervention and whether baseline falls self-efficacy affects changes in physical function. However, to our knowledge, no such investigation has been done. The primary purpose of this study was to reveal the effects of short-term exercise intervention on falls self-efficacy in Japanese community-dwelling older people. The exercise program that we applied in this study included strength and functional training. The second purpose of this study was to evaluate the relationships between the baseline falls self-efficacy of the subjects and their changes in physical function to obtain information on whether falls self-efficacy would alter the effect of exercise intervention.

METHODS

This study was a single-blinded randomized controlled trial with assessments before and after the intervention. The study was conducted in three institutions in Japan: the Tokyo Metropolitan Institute of Gerontology, which is located in Itabashi City in the Tokyo metropolitan area; the University of Kitasato, which is located in Sagamihara City in Kanagawa Prefecture; and Kanagawa University of Human Services, which is located in Yokosuka City in Kanagawa Prefecture. This study was approved by the ethics committees of the Tokyo Metropolitan Institute of Gerontology. The exercise intervention and health education classes were conducted in the gymnastic rooms or halls of these institutions.

Participants

The participants were residents of these three cities in Japan. They were recruited via advertisements in publications in these communities and through clubs for the elderly in these areas. The details of this study were explained before the study began, and written informed consent was obtained from all participants. The inclusion criteria were as follows: community-dwelling, aged 65 and older, people who were ambulatory with or without assisting devices, and who did not meet any of the exclusion criteria. The exclusion criteria for the study were (1) cerebrovascular or cardiovascular accidents reported within the past 6 mos; (2) acute liver problems or the active phase of chronic hepatitis; (3) diabetes mellitus with a history of hypoglycemic attack, or with fasting levels of plasma glucose concentrations of 200 mg/dl or higher, or with complications such as retinopathy or nephropathy; (4) systolic blood pressure above 180 mm Hg or diastolic blood pressure above 110 mm Hg at rest; (5) diagnosis of severe heart disease or an acute orthopedic problem; (6) diagnosis of dementia or depression made by a medical doctor, or an inability to understand and follow the instructions of the research staff; and (7) restriction of physical activities by a medical doctor.

Of the 190 participants in the first investigation, 171 participants met the eligibility criteria. Twelve subjects who met the exclusion criteria were excluded (Fig. 1). Nine subjects were excluded by criterion 5 (four subjects for heart disease, four for orthopedic problems, and one for both heart disease and an orthopedic problem), one was excluded by criterion 3, one was excluded by criterion...
4, and one was excluded by criterion 6. The average age of the participants was 74.1 yrs. Participants were randomly assigned into the exercise intervention group or the health education group, which was the control group (Fig. 1). There were no differences between the two groups in any measurements after the randomization (Table 1).

First Investigation of the Participants

The first investigation consisted of prescreening according to inclusion criteria and baseline testing. For prescreening, we interviewed the participants about their medical conditions. In baseline testing, demographic and morphological factors (age, height, weight, body mass index), Tokyo

### TABLE 1 Time course of measurements

<table>
<thead>
<tr>
<th></th>
<th>Exercise Group</th>
<th>Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre</td>
<td>Post</td>
</tr>
<tr>
<td></td>
<td>n  Mean ± SD</td>
<td>n  Mean ± SD</td>
</tr>
<tr>
<td>Age</td>
<td>85 73.9 ± 5.0</td>
<td>86 74.4 ± 6.2</td>
</tr>
<tr>
<td>Height, cm</td>
<td>85 154.0 ± 9.1</td>
<td>86 59.2 ± 10.1</td>
</tr>
<tr>
<td>Weight, kg</td>
<td>85 56.3 ± 9.0</td>
<td>86 24.3 ± 3.2</td>
</tr>
<tr>
<td>BMI</td>
<td>85 23.8 ± 3.4</td>
<td>86 27.8 ± 1.8</td>
</tr>
<tr>
<td>MMSE (/30)</td>
<td>85 80.0 ± 1.0</td>
<td>86 103.4 ± 10.6</td>
</tr>
<tr>
<td>FES (10–40)</td>
<td>71 38.7 ± 2.2</td>
<td>65 38.7 ± 2.5</td>
</tr>
<tr>
<td>One-leg standing with eyes open, secs</td>
<td>72 37.8 ± 23.0</td>
<td>66 37.6 ± 22.5</td>
</tr>
<tr>
<td>One-leg standing with eyes closed, secs</td>
<td>72 6.0 ± 6.2</td>
<td>65 4.1 ± 3.4</td>
</tr>
<tr>
<td>Functional reach test, cm</td>
<td>72 35.1 ± 6.1</td>
<td>65 34.3 ± 5.6</td>
</tr>
<tr>
<td>Modified timed up-and-go test, secs</td>
<td>71 5.6 ± 2.0</td>
<td>65 5.6 ± 1.3</td>
</tr>
<tr>
<td>Preferred walking velocity, m/min</td>
<td>71 76.7 ± 14.4</td>
<td>66 74.2 ± 12.5</td>
</tr>
<tr>
<td>Maximum walking velocity, m/min</td>
<td>68 114.3 ± 21.8</td>
<td>65 132.0 ± 21.5</td>
</tr>
<tr>
<td>Sit and reach, cm</td>
<td>71 32.0 ± 9.8</td>
<td>65 32.6 ± 9.3</td>
</tr>
<tr>
<td>Handgrip strength, kg</td>
<td>71 28.5 ± 7.8</td>
<td>65 28.2 ± 7.1</td>
</tr>
<tr>
<td>Knee extension strength, N</td>
<td>71 320.3 ± 124.5</td>
<td>65 327.0 ± 104.0</td>
</tr>
<tr>
<td>TMIG index (/13)</td>
<td>71 12.3 ± 1.3</td>
<td>65 12.5 ± 0.8</td>
</tr>
</tbody>
</table>

FES, Falls Efficacy Scale; BMI, body mass index.
* Significantly greater than pre ($P < 0.01$).
† Significantly greater than pre ($P < 0.05$).
‡ Significantly smaller than pre ($P < 0.01$).
§ Significantly smaller than pre ($P < 0.05$).
Falls Self-Efficacy

We used the falls efficacy scale (FES) for the Japanese population that was derived from Tinetti’s FES.\textsuperscript{23} Tinetti’s FES was based on Bandura’s theory of self-efficacy. This original FES assesses a subject’s confidence in performing (without falling) each of 10 activities that are considered essential to independent living, including cleaning the house, getting dressed and undressed, preparing simple meals, taking a bath or shower, simple shopping, getting in or out of a chair, going up and down stairs, walking around the neighborhood, reaching into cabinets or closets, and hurrying to answer the phone. Although the scoring of the degree of confidence in accomplishing tasks was based on a scale from 1 to 10 in the original FES, we used a four-degree scale in the FES ranging from 1 (not at all sure) to 4 (very sure) because it was easier for the participants to use. The range of scores in this study was from 10 to 40, with a higher score revealing greater confidence. Tinetti et al.\textsuperscript{10} and Tennstedt et al.\textsuperscript{24} used an FES with four degrees of response in their research in 1998. Tennstedt et al.\textsuperscript{24} reported the Cronbach’s alpha of the measurement to be 0.90–0.93. In addition, the validity of this methodology for measuring falls self-efficacy in the Japanese population has been investigated, and Cronbach’s alpha in these studies ranged from 0.80 to 0.94 (these results were published in journals in Japanese).

Measurements of Physical Function

The physical functions described in the following paragraphs were measured before and after investigation. One-leg standing with eyes open and closed was measured as an indicator of static balance.\textsuperscript{25,26} The subjects performed the posture twice, and the measurement recorded was the maximum value of these two. The maximum value was set at 60 secs. The functional reach test was performed to measure dynamic balance.\textsuperscript{27,28} They performed the test twice, and the measurement recorded was the maximum value. A timed up-and-go test (TUG) was performed to measure functional balance.\textsuperscript{29} To represent maximal abilities to perform the task of the TUG test, we slightly modified the methodology of this test. The instruction to subjects was modified as follows: “Please return to the chair as quickly as you can without falling.” The subject performed the test twice. Then, we used the minimum time needed to carry out the task as TUG score. Preferred and maximum walking velocity were measured to represent walking abilities.\textsuperscript{30} Subjects were asked to walk a track that had a total length of 16 m. The examiner recorded the time that the subject took to walk along the middle 10 m of the track with a stopwatch, and the velocity was then calculated. The subjects tried the test twice, and the largest value (m/min) was adopted as the record. The sit-and-reach test was measured as an indicator of the flexibility of the hamstrings.\textsuperscript{31} The equipment used in this test was made by Takei Scientific Instruments Co., Ltd. Japan. Handgrip strength\textsuperscript{32} and isometric knee extension strength\textsuperscript{16,33} were measured to represent upper- and lower–dominant limb strength. Handgrip strength was measured using a handheld dynamometer made by Takei Scientific Instruments Co., Ltd. Japan. Isometric knee extension strength was measured while subjects were sitting on a treatment table with their knees and hips at 90 degrees of flexion. We used a handheld dynamometer made by OG Giken Co., Ltd. Japan. The dominant leg was measured twice, and the recorded measurement was the maximum value.

We applied these measurements because of their reliability and convenience and because they have frequently been used in research studies to show physical frailty and the relationship of the balance function to the occurrence of falls. Because we were thinking about how to disseminate the preventive approach we were studying, we decided that it was necessary to simplify our methodology and use low-cost equipment. These assessments were performed by trained physical therapists and research assistants. They were blinded to the group allocation of the participants.

Exercise Intervention Protocol

Exercise intervention was conducted at a gymnastic room in each of three institutions. We used a training protocol that included progressive resistance training and balance training. This exercise program was constructed according to the American College of Sports Medicine guidelines\textsuperscript{34} and other research.\textsuperscript{35,36} The exercise was conducted by an interdisciplinary team that consisted of medical and fitness staff. The medical staff performed medical management, developed appropriate emergency response plans, and trained their staff.

The characteristics of this program are listed below: (1) the main objective was the improvement of physical function in older people; (2) the physical functions focused on in this program were...
muscular strength of the lower extremities, balance functions, flexibility, and daily functions such as climbing stairs; (3) instruction was given by an interdisciplinary team composed of a physician or nurse, a physical therapist, and a sports instructor, with the physical therapist managing the programs of the participants with pain; (4) the program was a group program in which instruction was given to fewer than ten subjects; (5) the duration of the program was 3 mos, with classes conducted for 1.5 hrs, twice a week; and (6) there were three periods in the program: the conditioning period, the muscular strength–enhancement period, and the functional training period. The target and intensity of training was individually set in each period.

Every session began with a 10- to 15-min warm-up consisting of gentle stretching and light exercise and ended with a 10-min cool-down of stretching. During the conditioning period, we focused primarily on physical conditioning. The subjects spent a great deal of time doing exercises that stretched their muscles. The subjects were also familiarized with weight-training machines and the technique of training at low intensity with frequent repetitions. At the beginning of the muscular strength–enhancement period, the subject’s one-repetition maximum was measured directly if the subject could perform machine training in an adequate manner. One-repetition maximum was defined as the maximum weight that could be lifted through a full range of motion with proper form. The subjects were trained using four weight-training machines (leg press, leg extension, hip abduction, and rowing; made by Tanren Co., Ltd. Japan). At the beginning of the strengthening phase, the intensities were set individually at 60% or more of their one-repetition maximum for two or three sets of ten repetitions. To avoid joint pain and muscle soreness, we applied these intensities at the beginning of the strength-enhancement period. We believed that such uncomfortable feelings would lead subjects to dropout. The participants who could not lift weight in an adequate manner continued conditioning training. If there were participants with specific physical problems such as joint pain or immobilization, the physical therapist treated their problems individually. The resistance was increased if the subjects were able to effortlessly complete three sets of ten repetitions. A rest of about 2 min or more was given between the sets of machine training. The subjects also underwent balance training in this period. During the functional training period, in addition to the high-intensity balance training described above, the program included intensive balance training and functional training exercises. The tasks of balance and functional training progressively increased in difficulty on the basis of set criteria and depending on an individual’s ability. Functional training that required more dynamic and rapid movement of the center of gravity was also introduced. For example, the target of the training at the early phase of the training period was to stand stably against small perturbations with a narrow base of support. During the first phase of the training period, correct trunk- and lower-body alignment in the standing position with wide lower-extremity stances was emphasized. The base of support was gradually reduced as their lower-extremity stances narrowed. The subjects who could not stand stably with correct alignment were allowed to use simple stabilizing equipment such as a chair or parallel bars. If we wanted to increase the difficulty of the static balance training, subjects were asked to stand on one foot or to raise their heels or toes up while standing without stabilizing equipment. We did not use specialized devices to create perturbations. The small perturbations that were used in this study were introduced by having subjects raise their arms by themselves, twist their trunk, catch or throw an object, be pushed by the therapist, and so on. Additionally, a soft foam surface that made it difficult to control the sway of their body mass was introduced to their base of support. In the next phase, the subjects tried to move their center of gravity widely and rapidly in their base of support with coordination of the lower extremities and trunk. In the final phase, the subjects performed high-level functional tasks such as side walking while crossing their steps, walking with bending their knees, jumping and landing on the floor, and so on.

Health Education Program

The subjects in the health education group, which functioned as a control group, received 1.5 hrs of lectures on health promotion for older people twice a month for 3 mos. The contents of the lectures were intended to help older people age successfully. The titles of the lectures in the health education programs included “Conditions for Successful Aging,” “Aging and Cognitive Function,” “Fall Prevention for Seniors,” “Health of Vessels,” “Gait Pattern of Seniors,” and “Knowledge of Resistance Training.” Lecture of the same titles were performed in each of the three institutions.

Statistical Analyses

After the randomization, we evaluated the difference in each measurement between the two groups using the unpaired t test. The improvement of measurements was evaluated by the paired t test and the Wilcoxon rank test. Because many of the subjects had perfect FES scores in this research, we dichotomized the subjects into two subgroups, one for those whose FES score was 40 and one for those
whose score was less than 40, and then we performed one-way analysis of variance to evaluate the differences between the groups in the magnitude of the change in physical function. To evaluate the relationships between the baseline FES score and physical functional measurements at baseline and the change in these measurements after the intervention, we used the Spearman rank-correlation coefficient. We hypothesized that the subjects with lower FES scores at baseline could improve their physical function more than subjects with higher FES scores. The magnitude of the changes in each measurement was calculated as the postinvestigation value minus the baseline investigation value. (One exception to this was the change in the TUG, which was calculated as the baseline value minus the postinvestigation value.) Stepwise linear regression analysis was applied to evaluate the associations between the magnitude of the changes in the measurements and the FES score. The dependent values were the magnitudes of the change in measurements during the intervention period, and the independent values were the baseline FES and physical functional measurements at baseline as well as group affiliation (exercise or education). We used SPSS 13.0J for Windows statistical software (SPSS Inc., Chicago, IL). P values less than 0.05 were considered statistically significant.

RESULTS

Twelve men and 22 women were not able to complete the intervention. Except in knee extensor strength, there were no differences in the baseline measurements between the subjects who completed the intervention and those who did not. Fourteen of the subjects who dropped out were from the exercise intervention group, and 20 were from the control group. In those subjects who did not complete the intervention, there were also no differences in baseline measurements between the groups. The reasons for not completing the intervention included loss of interest or an accident that did not originate directly with these programs, including hospitalization and bad health conditions such as colds. Missing values were found in three subjects’ FES data (two at baseline, and another at postinvestigation), and these data were excluded from the analysis.

There were no improvements in the FES score in either subgroup (Table 1). Also, there was no difference between the exercise intervention group and the control group in terms of FES score, both before and after the intervention period. In the exercise group, TUG, sit and reach, and TMIG index improved significantly, but handgrip strength decreased significantly (Table 1). In the control group, one-leg standing with eyes closed and sit-and-reach score improved significantly, but MWS, handgrip strength, and knee extension strength decreased significantly (Table 1). The results of one-way analysis of variance are displayed in Table 2. There was a significant difference among the four groups in the magnitude of the change in the maximum walking velocity ($P < 0.004$), and $P$ value was almost significant in knee extension strength ($P = 0.05$), but no difference was seen in the other measurements (Table 2). Table 3 shows the correlation coefficient between the FES score at baseline and the magnitude of the change in physical measurements. There were significant correlations between the FES score at baseline and the change in maximum walking velocity ($r = -0.268, P = 0.018$) and knee extension strength ($r = -0.248, P = 0.040$) in the exercise group. The results of linear regression analysis (stepwise) are displayed in Table 4. The baseline FES score was extracted as an element to explain the change in one-leg standing with eyes open.

DISCUSSION

The purposes of this study were to examine the effects of short-term exercise intervention including strength, balance, and functional training on FES and to reveal the associations between baseline FES and the changes in physical function in an elderly Japanese population.

The dropout rate in this study was thought to be low, and no adverse event was reported during the intervention period. However, there were no improvements in FES score in either group. There have been some studies that have attempted to reduce the fear of falling and to improve the FES. The results of these trials were mixed. Hornbrook et al. failed to reduce the fear of falling. Tinetti et al. demonstrated that multidimensional intervention reduced the fear of falling and increased fall efficacy. Maki et al. pointed out that programs that reduce the risk of falling may not always reduce the fear of falling, because this fear is, to some extent, independent of the risk of falling. According to previous studies, exercise interventions improved falls self-efficacy, but so did education programs. Social support and communication about falling are important covariates of fear of falling and can provide effective strategies to improve mental well-being.

In the current randomized study, neither group improved FES. The subjects of this study had fairly high physical function, and about half of these subjects had perfect scores on the FES. Their ability to improve their physical fitness and FES might have already reached its maximum level, so the results we obtained may indicate a ceiling effect. Therefore, we could not conclude that these short-term interventions were ineffective for improving FES in the community-dwelling older population. Additionally, it is difficult for us to explain why both groups improved
TABLE 2 Magnitude of the change in measurements in four subgroups (results of one-way analysis of variance)

<table>
<thead>
<tr>
<th>Measurements</th>
<th>Exercise Group</th>
<th>Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FES = 40</td>
<td>FES &lt; 40</td>
</tr>
<tr>
<td></td>
<td>n  Mean ± SD</td>
<td>n  Mean ± SD</td>
</tr>
<tr>
<td>One-leg standing with eyes open, secs</td>
<td>41  3.46 ± 19.21</td>
<td>30  1.39 ± 20.93</td>
</tr>
<tr>
<td>One-leg standing with eyes closed, secs</td>
<td>41  0.79 ± 6.58</td>
<td>30  0.09 ± 6.85</td>
</tr>
<tr>
<td>Functional reach test, cm</td>
<td>41  -1.91 ± 5.05</td>
<td>30  0.46 ± 6.22</td>
</tr>
<tr>
<td>Modified timed up-and-go test, secs</td>
<td>41  0.30 ± 1.05</td>
<td>30  0.22 ± 0.86</td>
</tr>
<tr>
<td>Preferred walking velocity, m/min</td>
<td>41  -0.94 ± 10.36</td>
<td>29  1.07 ± 11.99</td>
</tr>
<tr>
<td>Maximum walking velocity, m/min</td>
<td>40  -12.5 ± 20.37</td>
<td>28  0.79 ± 18.81</td>
</tr>
<tr>
<td>Sit and reach, cm</td>
<td>40  1.24 ± 5.59</td>
<td>30  2.67 ± 6.08</td>
</tr>
<tr>
<td>Handgrip strength, kg</td>
<td>40  -1.66 ± 4.46</td>
<td>30  -1.36 ± 4.36</td>
</tr>
<tr>
<td>Knee extension strength, N</td>
<td>40  -18.25 ± 78.55</td>
<td>29  22.69 ± 88.50</td>
</tr>
<tr>
<td>TMIG index</td>
<td>40  0.18 ± 0.59</td>
<td>30  0.20 ± 0.85</td>
</tr>
</tbody>
</table>

FES, Falls Efficacy Scale; IADL, instrumental activities of daily living.
duration of the intervention might have been too short, or the training stimulus might have been too modest for vigorous older people to improve their self-efficacy and physical function. Mendes de Leon et al.\textsuperscript{9} reported in their prospective study that low self-efficacy (FES) was particularly predictive of functional decline among older individuals who showed a decline in physical performance at 18 mos of follow-up, but among those who did not decline in physical performance, the predictive value was minimal. It is possible that there might be a bias in the subjects of this study. Our data suggested that participants tended to be more vigorous than average persons in the community. Also, the evaluation methods in this study did not sufficiently reflect the subjects’ changes in physical function. A further study that includes

<table>
<thead>
<tr>
<th>TABLE 3 Correlation coefficient between FES score at baseline and change in physical measurements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Changes in Measurements</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>One-leg standing with eyes open, secs</td>
</tr>
<tr>
<td>One-leg standing with eyes closed, secs</td>
</tr>
<tr>
<td>Functional reach test, cm</td>
</tr>
<tr>
<td>Modified timed up-and-go test, secs</td>
</tr>
<tr>
<td>Preferred walking velocity, m/min</td>
</tr>
<tr>
<td>Maximum walking velocity, m/min</td>
</tr>
<tr>
<td>Sit and reach, cm</td>
</tr>
<tr>
<td>Handgrip strength, kg</td>
</tr>
<tr>
<td>Knee extension strength, N</td>
</tr>
<tr>
<td>TMIG index</td>
</tr>
</tbody>
</table>

FES, Falls Efficacy Scale.

<table>
<thead>
<tr>
<th>TABLE 4 Linear regression model (stepwise) for the changes in measurements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent Variables (Change in Measurement)</td>
</tr>
<tr>
<td>One-leg standing with eyes open, secs</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>One-leg standing with eyes closed, secs</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Maximum walking velocity, m/min</td>
</tr>
<tr>
<td>Sit and reach, cm</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

$\beta$, standard partial regression coefficient; FES, Falls Efficacy Scale.
more subjects with lower levels of FES and that examines the effects of longer exercise interventions might be necessary.

REFERENCES


February 2007 Falls Self-Efficacy 141

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.
Cardiovascular Disease in Spinal Cord Injury
An Overview of Prevalence, Risk, Evaluation, and Management

ABSTRACT

Cardiovascular disease is a growing concern for the spinal cord–injured (SCI) population. For long-term SCI, morbidity and mortality from cardiovascular causes now exceeds that caused by renal and pulmonary conditions, the primary causes of mortality in previous decades. Although risk estimates commonly used for ambulatory individuals have not been established from follow-up studies in SCI, nearly all risk factors tend to be more prevalent in SCI subjects compared with ambulatory subjects. These risks include a greater prevalence of obesity, lipid disorders, metabolic syndrome, and diabetes. Daily energy expenditure is significantly lower in SCI individuals, not only because of a lack of motor function, but also because of a lack of accessibility and fewer opportunities to engage in physical activity. Autonomic dysfunction caused by SCI is also associated with several conditions that contribute to heightened cardiovascular risk, including abnormalities in blood pressure, heart rate variability, arrhythmias, and a blunted cardiovascular response to exercise that can limit the capacity to perform physical activity. Thus, screening, recognition, and treatment of cardiovascular disease should be an essential component of managing individuals with SCI, and judicious treatment of risk factors can play an important role in minimizing the incidence of cardiovascular disease in these individuals. This article reviews the cardiovascular consequences of chronic SCI, including the prevalence of cardiovascular disease and risk factors unique to these individuals, and provides a synopsis of management of cardiovascular disease in this population.

Key Words: Cardiovascular Disease, Spinal Cord Injury, Renal, Pulmonary Diseases, Metabolic Syndrome, Exercise Testing, Mortality

Spinal cord injury (SCI) is a serious medical condition with considerable functional, psychological, and socioeconomic sequelae. Although there have been major advances in neurological treatment for SCI, the mortality rate for this condition remains high relative to ambulatory populations.1–3 Historically, respiratory and renal conditions have been the most prevalent comorbidities in...
the SCI population, and they remain important causes of mortality. However, data published in recent years suggest that cardiovascular disease (CVD) has emerged as the leading cause of mortality in chronic SCI. Morbidity from cardiovascular causes, particularly coronary artery disease (CAD), is high relative to ambulatory subjects, and CAD tends to occur earlier in SCI individuals than among ambulatory populations. A major contributor to the heightened risk of CVD in SCI is the fact that risk factors, including hyperlipidemia, obesity, and diabetes, have been shown to be comparatively high among individuals with SCI. The recognition and treatment of CVD is an emerging clinical challenge in this population.

An additional contributing factor to the high cardiovascular morbidity and mortality in SCI is the sedentary lifestyle and reduced physical function associated with loss of motor function. SCI is also characterized by a disruption of the normal autonomic cardiovascular control mechanisms, and there is growing recognition that this further contributes to cardiovascular risk. The latter occurs as a result of a variety of physiologic changes related to cardiovascular control that are observed in SCI, including loss of normal regulation of the peripheral vasculature, autonomic dysreflexia (AD), and a higher prevalence of cardiac rhythm disturbances. The major cardiovascular concerns associated with SCI are outlined in Table 1. The purpose of this review article is to provide an overview of the cardiovascular consequences of chronic SCI, including the prevalence of CVD and risk factors unique to these individuals, and a synopsis of management of CVD in this population.

**Prevalence of CVD in SCI**

Precise estimates of the prevalence of CVD in SCI are complicated by the comparatively high prevalence of latent heart disease and by the misclassification of CVD attributable to concomitant disorders. Nevertheless, studies are consistent in demonstrating a higher prevalence of CVD among SCI individuals compared with that in ambulatory populations. Groah et al. studied 545 SCI subjects surviving at least 25 yrs after injury and observed that the risk of developing CVD was associated with both the level and extent of injury. Tetraplegic level of injury was associated with a 16% higher risk of all CVD (CAD, hypertension, cerebrovascular disease, valvular disease, and dysrhythmias) and a fivefold increase in cerebrovascular disease, but paraplegic subjects had a 70% greater risk of CAD. Complete injury conferred a 44% greater risk of overall CVD. Bauman and colleagues reported that the prevalence of silent ischemia in a middle-aged (mean 52 yrs) cohort of paraplegic subjects was 65% based on nuclear imaging. Other studies have reported lesser, but nevertheless substantial, prevalence rates of asymptomatic CVD in SCI populations (ranging from approximately 25% to more than 50%). The prevalence rates of symptomatic CVD have similarly ranged from approximately 30% to more than 50%. In contrast, among age-matched able-bodied populations, the prevalence of CVD is typically reported to be in the range of 5–10%.

In terms of mortality from cardiovascular causes, cohorts of subjects with chronic SCI have been reported to have both higher cardiovascular mortality rates and mortality occurring at earlier ages compared with able-bodied subjects. This is particularly true among subjects with SCI of long duration. Whiteneck and colleagues reported that CVD was the leading cause of mortality in persons with SCI of more than 30-yr duration; approximately half of SCI subjects in this cohort died of cardiovascular causes. In an analysis of >28,000 SCIs occurring between 1973 and 1998 among subjects admitted to two SCI health care systems, heart disease was the leading cause of mortality after the first year of injury. Although morbidity and mortality rates have not been established with certainty, the available data suggest that major efforts need to be made to identify risk factors for CVD that are modifiable, and appropriate intervention strategies should be developed to reduce these rates of cardiovascular morbidity and mortality in persons with SCI.

**TABLE 1** Cardiovascular concerns in spinal cord injury

- Higher prevalence of cardiovascular disease
- Greater morbidity and mortality from cardiovascular causes
- Heightened cardiovascular risk factors: Low high-density lipoprotein cholesterol, High total cholesterol and low-density lipoprotein, Elevated C-reactive protein
- Higher prevalence of obesity and greater visceral adipose tissue
- Increased rate of smoking
- Physical inactivity
- Higher prevalence of insulin resistance, diabetes, and metabolic syndrome
- Blood pressure abnormalities (orthostatic hypotension, autonomic dysreflexia)
- Deep vein thrombosis, thromboembolic events
- Rhythm disturbances
- Bradycardiacs, particularly in the acute phase, (e.g., bradycardia, A-V block, cardiac arrest)
- Reduced heart rate variability
- Blunted cardiovascular response to exercise

February 2007

Cardiovascular Disease in Spinal Cord Injury 143
Risk Factors for CVD in SCI

A heightened prevalence of virtually all the major risk factors for CVD exists for persons with SCI, and this is one of the major challenges for clinicians who treat this condition. An algorithm for guiding CVD risk management, modified from the American Heart Association Guidelines for Primary Prevention of Cardiovascular Disease and Stroke, is presented in Table 2. Although algorithms for reducing CVD risk have not been designed specifically for SCI, these strategies can be generally used to guide risk reduction for those at risk for CVD.

Diabetes/Metabolic Syndrome

Diabetes and metabolic syndrome are two closely linked conditions associated with CVD risk and are considered by the American Heart Association to be major risk factors for heart disease. Metabolic syndrome is a prediabetic state that has been strongly linked to both heart disease and diabetes mellitus, its presence roughly doubles the risk of CVD mortality. The metabolic syndrome is closely associated with obesity (increased waist circumference), atherogenic dyslipidemia (high triglycerides; low high-density lipoprotein [HDL] cholesterol; increased small, dense low-density lipoprotein [LDL] cholesterol; and increased apolipoprotein B), increased blood pressure, insulin resistance (hyperinsulinemia, glucose intolerance, increased uric acid), a prothrombotic state (increased plasminogen activator inhibitor [PAI-1], increased blood viscosity, increased plasma fibrinogen), and proinflammatory conditions (increased C-reactive protein, or CRP). The prevalence of the metabolic syndrome and its individual components have been shown to be high among individuals with SCI. Lee and colleagues observed that the metabolic syndrome was present in 23% of SCI individuals (roughly double that of populations of similar age) and that prediabetes and cardiovascular risk scores by Framingham criteria were elevated relative to those reported among ambulatory populations.

The use of ambulatory guidelines for the determination of metabolic syndrome in the SCI population may not be appropriate in the context of the anthropometric and physiologic changes associated with chronic SCI. In particular, the use of waist circumference may not adequately reflect abdominal adipose tissue or abdominal visceral adipose tissue in SCI individuals, given the particularly sedentary nature of chronic SCI. In some studies, body mass index has been substituted for waist circumference; however, body mass index often underestimates body fat in individuals with SCI. Although it is assumed that the sedentary lifestyle imposed by chronic SCI increases central obesity, there is no current consensus for a clinically useful measure of obesity in the SCI population. Other criteria for determination of the metabolic syndrome have not been adequately defined for the SCI population. Although persons with SCI have generally been found to have greater lipid abnormalities compared with able-bodied individuals, it is not clear how this affects their prediabetic status. In addition, hypertension and glucose metabolism require further examination as components of the metabolic syndrome in persons with SCI.

Lipid Disorders in SCI

Abnormal lipid values have long been established as risk factors for the development of diabetes and heart disease. After SCI, there is a tendency toward elevated LDL cholesterol and total cholesterol as well as lower HDL cholesterol levels compared with able-bodied persons. Persons with tetraplegia tend to have a greater number of lipid abnormalities than their paraplegic counterparts. The greater degree of dyslipidemia found in the SCI population contributes significantly to their increased CVD risk. Abnormal lipids are generally modifiable with changes in physical activity and diet along with the use of statins; however, among persons with SCI, the degree of dyslipidemia is more strongly linked to the duration of injury than to diet. This suggests that the metabolic changes and physical inactivity associated with SCI may have significant consequences for the prevalence of dyslipidemia in this population.

Inflammatory Markers

Inflammation is increasingly recognized to have an important role in the development of CVD. The most widely studied inflammatory marker, CRP, has been found to be elevated in both acute and chronic SCI. Elevated CRP is observed in SCI individuals both with and without urinary tract infections, suggesting that it may be caused more by some underlying disease state than by the injury itself. To date, only one study has investigated the relationship between CRP and CVD risk in persons with SCI. In that study, CRP was significantly associated with the presence of other well-known CVD risk factors, including multiple lipid abnormalities, metabolic syndrome, insulin resistance, and elevated Framingham risk. The mean high-
### TABLE 2 Guide to primary prevention of cardiovascular diseases

<table>
<thead>
<tr>
<th>Risk Intervention</th>
<th>Recommendations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Smoking:</strong></td>
<td></td>
</tr>
<tr>
<td>Goal</td>
<td>Ask about smoking status as part of routine evaluation. Reinforce nonsmoking status.</td>
</tr>
<tr>
<td>Complete cessation</td>
<td>Strongly encourage patient and family to stop smoking.</td>
</tr>
<tr>
<td>Blood pressure control:</td>
<td>Provide counseling, nicotine replacement, and formal cessation programs as appropriate.</td>
</tr>
<tr>
<td>Goal</td>
<td>Measure blood pressure in all adults at least every 2.5 yrs.</td>
</tr>
<tr>
<td>&lt;140/90 mm Hg</td>
<td>Promote lifestyle modification: weight control, physical activity, moderation in alcohol intake, moderate sodium restriction. If blood pressure is around 140/190 mm Hg after 3 months of life-habit modification, or if initial blood pressure is greater than 160/100 mm Hg, add blood pressure medication; individualize therapy to patient's other requirements and characteristics.</td>
</tr>
<tr>
<td><strong>Lipid management</strong></td>
<td></td>
</tr>
<tr>
<td>Primary goal</td>
<td>Ask about dietary habits as part of routine evaluation.</td>
</tr>
<tr>
<td>LDL &lt;160 mg/dl if zero or one risk factors</td>
<td>Measure total and HDL cholesterol in all adults 20 yrs or older, and assess positive and negative risk factors at least every 5 yrs.</td>
</tr>
<tr>
<td>or LDL &lt;130 mg/dl if ≥2 risk factors</td>
<td>For all persons: promote AHA Step I diet (30% fat, &lt;10% saturated fat, &lt;300 mg/dl cholesterol), weight control, and physical activity.</td>
</tr>
<tr>
<td>Secondary goals</td>
<td>Measure LDL if total cholesterol is around 240 mg/dl or &gt;200 mg/dl with two or more risk factors or if HDL &lt;35 mg/dl</td>
</tr>
<tr>
<td>HDL &gt;35 mg/dl</td>
<td>Suggested drug therapy for high LDL levels (&gt;160 mg/dl) (drug selection priority modified according to TG level)</td>
</tr>
<tr>
<td>TG &lt;200 mg/dl</td>
<td>TG 200–400 mg/dl</td>
</tr>
<tr>
<td>Statin</td>
<td>Statin</td>
</tr>
<tr>
<td>Resin</td>
<td>Niacin</td>
</tr>
<tr>
<td>If LDL:</td>
<td></td>
</tr>
<tr>
<td>&gt;160 mg/dl with zero or one risk factor, or &gt;130 mg/dl on two occasions with more than two risk factors; then —Start Step II diet (&lt;30% fat, &lt;7% saturated fat, &lt;200 mg/dl cholesterol) and weight control. —Rule out secondary causes of high LDL (LFTs, TPTs, UA).</td>
<td></td>
</tr>
<tr>
<td>If LDL:</td>
<td>Consider adding drug therapy to diet therapy for LDL levels greater than those listed above that persist despite Step II diet.</td>
</tr>
<tr>
<td>&gt;160 mg/dl plus two risk factors; or &gt;190 mg/dl; or &gt;220 mg/dl in men &lt;35 yrs; or in premenopausal women; then —Consider combined drug therapy (niacin, fibrates, statin)</td>
<td></td>
</tr>
<tr>
<td>If LDL goal not achieved, consider combination drug therapy.</td>
<td>Risk factors: age (men &gt;45 yrs, women &gt;55 yrs or postmenopausal), hypertension, diabetes, smoking, HDL &lt;35 mg/dl, family history of CHD in first-degree relatives (in male relatives &lt;55 yrs, female relatives &lt;65 yrs). If HDL &gt;60 mg/dl, subtract one risk factor from the number of positive risk factors.</td>
</tr>
</tbody>
</table>

Emphasize weight management and physical activity, avoidance of cigarette smoking. Niacin raises HDL. Consider niacin if patient has more than two risk factors and high LDL (except patients with diabetes).
sensitivity CRP level in these subjects (2.37 ± 2.1 mg/liter) placed them in a high-risk group on the basis of quintiles established for ambulatory individuals. However, no follow-up studies exist that have established the role of CRP in the development of atherosclerosis and CVD in SCI.

**Physical Inactivity in SCI**

The reduced physical function associated with SCI underlies a greater sedentary lifestyle and lower energy expenditure. Using a variety of techniques to quantify daily energy expenditure, studies have shown that SCI individuals have lower resting metabolic rates and, on average, expend significantly less daily energy than ambulatory subjects. This underlies a higher proportion of fat mass and a greater prevalence of obesity, contributing to a variety of related metabolic abnormalities associated with inactivity, including insulin resistance, lower HDL levels, and greater susceptibility to vascular inflammation.

Chronic immobilization associated with SCI also leads to a number of skeletal-muscle metabolic and structural abnormalities. Specific alterations in morphologic and contractile properties of skeletal muscle with chronic SCI have been shown using electromyography studies, biopsy, and magnetic resonance imaging. These alterations include lower protein content, an increase in myosin heavy-chain isoforms, reduced fiber cross-sectional area, and reduced force and fatigue characteristics with functional electrical stimulation. The inability to ambulate the lower limbs also contributes to the risk of deep vein thrombosis, in addition to the above-mentioned heightened prevalence of insulin resistance and low HDL levels.

Recent studies have shown that many of the metabolic and skeletal-muscle abnormalities associated with SCI can be partially reversed by endurance training with upper-body arm ergometry, functional electrical-stimulation training of the lower limbs, or their combination. Swimming, supported treadmill ambulation, and other adapted modes of training have also been used. Peak VO₂ has been shown to increase to a degree that is similar to or slightly less than that of ambulatory subjects, for instance, on the order of 10–20% after varying periods of training. These increases seem to be inversely proportional to the level of injury. Regular exercise has also been shown to favorably affect lipid profiles in SCI.

**TABLE 2 Continued**

<table>
<thead>
<tr>
<th>Risk Intervention</th>
<th>Recommendations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Physical activity:</strong></td>
<td><strong>Goal</strong></td>
</tr>
<tr>
<td><strong>Weight management:</strong></td>
<td><strong>Goal</strong></td>
</tr>
</tbody>
</table>

Modified from Pearson et al. 23
muscle mass, and increase isometric strength and endurance. In some studies, functional electrical-stimulation training has also improved lower-limb circulation and vasodilatory capacity, body composition, and insulin resistance in subjects with varying levels of SCI. However, ready access to these sorts of training regimens is lacking in the general SCI population.

Cardiovascular Consequences of Autonomic Dysfunction

Autonomic nervous system dysfunction causes a disruption of normal cardiovascular homeostasis, which itself increases the risk of CVD, particularly in higher-level injuries. Cardiovascular problems directly associated with autonomic dysfunction that have been extensively described in SCI include loss of vasomotor control leading to orthostatic hypotension, AD, reflex bradycardia and, in extreme cases, cardiac arrest. Other factors that increase the risk for CVD that are attributable to loss of supraspinal control include reduced heart rate variability (HRV); attenuated cardiovascular responses to activity, including reduced cardiac contractility; and changes to the skin microcirculation.

Blood Pressure Abnormalities

Arterial blood pressure is typically chronically low in individuals with SCI because of a reduction in sympathetic nervous system activity below the level of injury. Hypotension at rest and, in particular, orthostatic hypotension, contribute to homeodynamic instability in SCI. Orthostatic hypotension is characterized by dizziness, weakness, blurred vision, and syncope when shifting from the supine to the upright sitting position. The normal hemodynamic response to standing is an increase in heart rate and contractility via activation of autonomic reflexes through the carotid and aortic baroreceptors. This causes vagal inhibition and sympathetic stimulation, resulting in an increase in blood pressure. Baroreflex control of vascular tone is often absent in SCI, particularly in higher-level injuries, among whom orthostatic hypotension is common.

AD is a far more serious hemodynamic consequence of SCI; in extreme cases, it can be life threatening. AD is characterized by sympathetic hyperactivity, causing severe vasoconstriction and hypertension below the level of the lesion. Cerebrovascular accidents secondary to AD have been described and are thought to be an important cause of mortality in the SCI population. AD has been estimated to occur between 48 and 90% of individuals who are injured at level T6 and above. With AD, vasomotor reflexes above the level of injury attempt to lower blood pressure by increasing parasympathetic stimulation to the heart via the vagus nerve, which results in vasodilation, light-headedness, profuse sweating, and skin flushing. Because AD can trigger severe cardiovascular reactions, it represents a medical emergency that requires immediate treatment. Cardiovascular consequences of AD can include periods of severe bradycardia or tachycardia, a hypertensive emergency (severe, sustained hypertension causing significant organ damage or impairment), left ventricular failure, myocardial ischemia, or serious rhythm disturbances. Treatment requires immediate removal of the precipitating stimuli (most commonly, bladder distension) and pharmacologic stabilization of blood pressure.

HRV

HRV describes the quantification of beat-to-beat variations of the R-R interval on the electrocardiogram, measured over a period of time ranging from a few minutes to 24 hrs. Its recent appeal can be attributed to the fact that it is simple to measure, is noninvasive, and is known to mirror sympathetic and parasympathetic nervous system balance. Numerous studies published in the last 10 yrs have documented that abnormal (reduced) HRV strongly predicts risk for cardiac events. A great deal of research in recent years has employed HRV indices to assess autonomic balance in various chronic disease states, including neuromuscular disorders, neuropathy caused by diabetes, and SCI.

HRV indices to assess autonomic balance in various chronic disease states, including neuromuscular disorders, neuropathy caused by diabetes, and SCI. SCI provides a good model for the application of HRV, because the interruption of efferent sympathetic pathways innervating the cardiovascular system disrupts the normal autonomic nervous system balance, which is reflected in altered HRV patterns. Several groups have observed an inverse association between the level and completeness of injury and autonomic dysfunction expressed using HRV. These HRV findings reflect both parasympathetic downregulation and lessened sympathetic nervous input to the heart. In addition, they confirm that HRV is abnormal in persons with SCI relative to able-bodied individuals. Patterns of altered HRV have been suggested as useful in characterizing the physiology associated with injury level, and although this requires further study, these patterns may have diagnostic, prognostic, and therapeutic significance.

Cardiac Arrhythmias

High-level injuries are particularly prone to cardiac rhythm disturbances, which are most notable in the acute phase of injury. Studies suggest that this is largely attributable to heightened sympathetic tone. Arrhythmias after SCI can range from benign to fatal. Bradyarrhythmias are...
common in the acute phase, including AV block; in fact, bradycardia was shown to occur in the vast majority of consecutive patients referred to a neurologic intensive care unit after SCI. Lehmann and colleagues observed that 16% of cervical-injured patients experienced cardiac arrest during the 14 days after hospital admission for acute injury. Rhythm disturbances are less of an issue among patients with chronic injuries. Marcus and colleagues and Prakash et al. observed that the occurrence of premature ventricular contractions on routine resting electrocardiograms was similar between ambulatory individuals and those with chronic SCI. A consistent finding, however, has been a higher prevalence of ST elevation in chronic SCI compared with ambulatory subjects, suggesting a shift in autonomic balance favoring vagal tone.

**Blunted Cardiovascular Responses to Activity**

The magnitude of the physiologic response to exercise is diminished in SCI. Presumably, the attenuated cardiovascular and metabolic responses to activity reduce the well-known gains achieved by ambulatory subjects when exposed to regular physical activity. Many studies have reported a strong inverse association between the level of injury and peak VO2 achieved in SCI. Higher-level injuries generally prevent adequate sympathetic drive to increase heart rate beyond approximately 120–125 bpm. However, even those with injuries below the level of sympathetic outflow (T6) have lower stroke volumes at rest and attenuated cardiac output responses to exercise.

Loss of sympathetic tone also causes reduced myocardial preload and myocardial contractility, resulting in reduced stroke volume and cardiac output via the Frank–Starling mechanism. Chronically, these changes can lead to myocardial atrophy, although this finding is not universal.

The redistribution of blood flow to the active muscles during exercise that normally occurs in ambulatory individuals is largely absent after SCI, and, combined with the absence of intermittent contraction and relaxation of the skeletal muscles and absent or insufficient venoconstriction, venous return is inadequate, further blunting cardiac output. Blood tends to pool in the lower extremities, and compensatory increases in heart rate at rest and during exercise attempt to increase cardiac output in the presence of reduced ventricular volumes. Thus, arm ergometry exercise frequently results in hypotension because the metabolic demands of exercise are not matched by normal changes in peripheral vascular resistance. In addition, there is evidence that cardiac output may be lessened by intrinsically reduced cardiac contractility in SCI. Ventilation is significantly impaired in tetraplegia because of paralysis of the intercostal and abdominal musculature, reduced pulmonary compliance, reduced diaphragmatic excursion, and blunted chemoreceptor stimulation. This reduces inspiratory and expiratory pressures and all pulmonary function indices. In combination, these factors contribute significantly to a reduced capacity to adapt appropriately to a bout of exercise, resulting in early fatigue, general avoidance of exertion, and deconditioning, which further contribute to CVD risk.

**Considerations for Treatment and Evaluation of CVD in SCI**

In addition to the high prevalence of risk factors and their management, major cardiovascular considerations unique to chronic SCI include the detection and management of CAD, management of blood pressure (orthostatic hypotension and AD), and deep vein thrombosis. Although it has been demonstrated that SCI individuals are at higher risk for CVD (e.g., higher prevalence of lipid disorders, insulin resistance, and obesity, and reduced physical activity patterns), the extent to which multivariate models commonly used to estimate risk in ambulatory persons apply to SCI are unknown. Because of physiologic changes associated with SCI, the presence of SCI itself confers a higher risk for CVD. These changes include impaired autonomic responses that underlie abnormalities in resting heart rate, reduced HRV, reduced contractility, increased plasma rennin activity, and either heightened or blunted catecholamine production, depending on the level of injury and activity. Although the latter factors are not usually the first issues considered when assessing risk for CVD, each can portend a higher risk for cardiac events.

**Detection and Management of CAD**

Diagnosis and management of CAD in SCI is complicated by the high prevalence of asymptomatic disease, known as silent ischemia. Symptoms of CAD are frequently masked by the interruption of ascending afferent pain fibers in SCI; this presents a challenge in that lack of pain perception may cause CAD to be undetected and, therefore, untreated. Although silent ischemia is a concern in SCI, particularly in those with diabetes, the response to chest pain in general is, nevertheless, what should guide decisions regarding diagnostic studies. Traditional risk-factor scores used for ambulatory patients will likely predict a high probability for CAD in many SCI individuals, but the extent to which these scores relate to outcomes in SCI has not been explored. Thus, in general, careful treatment of risk factors is initially indicated rather than diagnostic testing.
The fact that persons with SCI are less likely to exert themselves to a level that would elicit angina also contributes to the underdiagnosis of CAD. In addition, diagnosis of CAD may be more difficult in SCI because exercise testing using arm ergometry is less likely to raise myocardial oxygen demand adequately to induce ischemic changes. Although arm ergometry has been shown to provide modest test characteristics in ambulatory patients (sensitivity of approximately 50%), there are no studies specifically addressing this issue in large groups of SCI patients referred for exercise testing for clinical reasons. Therefore, the initial diagnostic test is most commonly pharmacologic stress, along with an echocardiographic or nuclear evaluation. These diagnostic tests may also be indicated for preoperative assessment in high-risk patients.

Treatment of CVD Risk Factors

Comprehensive CVD risk-factor evaluation should be an integral part of every clinical visit for individuals with SCI. Clinicians need to be cognizant of both the high prevalence of CVD risk markers as well as risks that are unique to this population, and they should treat these risks judiciously. Because regular physical activity has been shown to improve lipid profiles and other risk factors in SCI individuals, physical activity should be strongly encouraged. Because of the physical and environmental limitations to physical activity participation for most persons with SCI, evaluation and follow-up by a physical therapist or exercise physiologist may help to optimize strategies to improve physical activity participation. Similarly, nutritional needs for persons with SCI are highly individualized because of wide differences in resting and daily energy expenditure. Routine consultation with a nutritionist may be helpful in the management of body weight and lipid profiles. This may be particularly important because dietary patterns of SCI individuals are generally poor.

Deep Vein Thrombosis

Deep vein thrombosis (DVT) can occur frequently during either the acute or chronic phase of SCI. DVT is an important clinical concern throughout the rehabilitation period because mortality from pulmonary embolus is increased the year after injury. Even in the presence of anticoagulant therapy, the incidence of DVT in SCI has been shown to range from 7 to 100%, depending on age and severity of injury, and 2.7% develop fatal pulmonary embolism. Treatment focuses on prevention of pulmonary embolism and prevention of recurrent DVT. After SCI, prophylactic treatment is usually recommended for 3 mos. Once the diagnosis of DVT is made, the usual treatment regimen includes low–molecular weight heparin initially for treatment of the thrombus, and prophylactic warfarin for at least 6 mos. The International Normalized Ratio (INR) should be maintained between 2 and 3.

Arrhythmias

Absence of sympathetic tone underlies a number of different rhythm disturbances, particularly in the acute phase of SCI. These can include severe bradycardia, A-V block, and cardiac arrest. Treatment for these often requires atropine administration, but in rare instances, pacemaker implantation may be required. In chronic SCI, arrhythmia management should be guided by the presence or absence of CVD, similar to ambulatory individuals. Presence of structural heart disease and poor left ventricular function are the most important factors in determining both prognosis and whether pharmacologic treatment is warranted. In the absence of CVD, drug therapy is not usually recommended for asymptomatic, isolated premature ventricular contractions, even when they are frequent. In the presence of CVD or arrhythmias causing symptoms, however, treatment initially involves beta blockade, with more serious rhythm disturbances requiring antiarrhythmic drugs, radiofrequency ablation, or, with life-threatening arrhythmias, an automatic implantable defibrillator.

Summary

Abnormalities of the cardiovascular system have increasingly become a concern for treating individuals with SCI. CVD is now recognized as a leading cause of morbidity and mortality, and autonomic nervous system dysfunction underlies several cardiovascular irregularities that contribute to CVD. These include an accelerated decline in cardiovascular function with aging, and a heightened prevalence of virtually all the major risk factors for CVD. In the absence of follow-up trials specifically addressing the role of clinical and lifestyle factors on the incidence of CVD in SCI, algorithms for primary prevention developed for ambulatory individuals are applicable to SCI (Table 2). Better recognition of the importance of CVD in SCI, which is often asymptomatic and thus undertreated, has the potential to significantly reduce morbidity and mortality from CVD. Judicious treatment of lipids and insulin resistance and promotion of creative ways to enhance physical activity are important treatment components for this condition and will help temper the incidence of CVD for persons with SCI.

REFERENCES


February 2007

Cardiovascular Disease in Spinal Cord Injury

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.
Phrenic Nerve Stimulation in the Evaluation of Ventilator-Dependent Individuals with C4- and C5-Level Spinal Cord Injury

ABSTRACT


Three individuals with C4 or C5 spinal cord injuries (SCI) were seen in follow-up for management of their late complications, which included impaired ventilation. Electrodiagnostic studies were performed on all three as part of the assessment of the function of their phrenic nerves and diaphragm muscles in relation to their need for mechanical ventilator support. Each patient had evidence of lower–motor neuron injury to the phrenic nerves. Two of the patients who initially displayed small-amplitude (<0.1 mV) compound muscle action potentials (CMAP) bilaterally were later reevaluated during the course of their observation in the outpatient rehabilitation clinic. The CMAP amplitude of the diaphragm increased in these two cases during the 3–11 mos after SCI. Evidence of nerve recovery occurred in parallel with improvements in pulmonary function testing and was followed by successful weaning from the ventilator. These individuals both gained ventilator independence after the CMAP amplitude of least one hemidiaphragm was >0.4 mV. In the third case, early failure of ventilator weaning was reported to the patient as a poor prognostic sign. At the time of our first evaluation 11 mos after injury, a CMAP of 1.0 mV was seen on the right, with an absent response on the left. In case 3, the needle electromyogram demonstrated voluntary active motor unit action potentials that provided additional electrophysiologic support for phrenic nerve function. Phrenic nerve–conduction studies can provide useful measures in assessing the recovery of lower–motor neuron diaphragm function in relation to impaired ventilation in individuals with C4- or C5-level SCI.

Key Words:  Spinal Cord Injury, Mechanical Ventilation, Phrenic Nerve, Nerve-Conduction Testing, Electromyography, Rehabilitation
Diaphragm paralysis is a devastating and functionally limiting component of high and midlevel cervical spinal cord injuries (SCI) that frequently results in prolonged ventilator dependency. Appropriate management of the ventilator-dependent patient requires attention to all aspects of breathing, including muscles of inspiration, muscles of expiration, aspiration, clearing of bronchial secretions, injury of the lung parenchyma, bronchial constriction, and proper nutrition. Previous authors have employed phrenic nerve–conduction testing as a part of the evaluation of individuals with SCI and ventilatory failure. Assessment of phrenic peripheral nerve function is useful in planning for the use of phrenic-stimulation pacing for activation of the diaphragm. Studies have suggested repeating this testing in a serial manner. The compound muscle action potentials (CMAPs) change in amplitude over time in different subjects. This suggests that phrenic nerve–conduction studies have potential as a clinical indicator for recovery of diaphragm function in individuals with C4- or C5-level SCI and as a possible adjunct to conventional assessments such as pulmonary function tests.

CASE 1

A 33-yr-old Caucasian male sustained a C4 burst fracture after a diving accident. He underwent surgical decompression and anterior cervical fusion. His physical examination findings were consistent with a C4 sensory- and motor-level SCI, American Spinal Injury Association (ASIA) class A, with a zone of partial preservation of sensation to the level of T10. He was ventilator dependent throughout his hospital course, which extended beyond 2 mos.

An electrodiagnostic study was requested and was performed 3 wks after his initial injury, with the referring physician requesting a determination of whether the might become a suitable candidate for diaphragmatic pacing. Each phrenic nerve was stimulated in the neck near the sternocleidomastoid muscle, with ipsilateral surface electrode recording conducted over the diaphragm using the methods described by Bolton and Markand et al. The active (E1) electrode is placed on the xyphoid process, and the reference electrode (E2) is placed on the lower costal margin at the anterior axillary line. Monopolar needle insertion for electromyography was performed using Koepke’s method at the anterior axillary line in the most inferior intercostal space (between the ninth and tenth ribs), and a surface reference electrode was employed. The needle electromyogram (EMG) of the diaphragm displayed 2–3+ positive sharp waves and fibrillations, and there were no voluntary motor unit action potentials (MUAP). The diaphragm CMAPs were absent at that time. On the basis of this evidence of peripheral motor axon loss, the patient was not considered a candidate for phrenic nerve–stimulation pacing at that early stage.

The patient remained ventilator dependent, with several subsequent, failed attempts at weaning. A repeat electrodiagnostic study was performed 1 mo later. At that time, the EMG showed a small number of increased duration, polyphasic MUAPs firing at a rate of greater than 20 Hz, and 1–2+ positive sharp waves and fibrillations. The phrenic nerve CMAPs were reproducibly recorded with amplitudes of 0.03 and 0.04 mV (Table 1).

After discharge to an extended care facility on a mechanical ventilator, the patient was reevaluated 4 mos postinjury after readmission to the acute care hospital. The phrenic nerve–conduction studies were repeated 4 mos after injury. The CMAP amplitudes were significantly improved and were 0.49 mV on the right and 0.13 mV on the left. Atelectasis of the left lower lobe with otherwise normal lung fields was seen on plain radiographs, and diminished, active diaphragm motion was seen on fluoroscopy. Recovery of inspiratory function was confirmed with pulmonary function testing (results not available), and weaning efforts were initiated. Unassisted breathing was successfully achieved after 3 days. The patient has remained off the ventilator since then for more than a year. There was no improvement in his neurological level during this time (i.e., he remained C4, ASIA A, sensory and motor).

CASE 2

A 62-yr-old Caucasian male sustained a C4–5 subluxation after a fall. He underwent surgical spinal discectomy, decompression, and fusion. His neurological examination findings were compatible with a C5, sensory and motor, ASIA A SCI. He was ventilator dependent when seen for evaluation of phrenic nerve function by electrodiagnostic studies 4 wks after his injury. His diaphragm CMAP amplitudes were less than 0.15 mV bilaterally, and he displayed positive waves and fibrillations and no clear diaphragmatic activity on EMG. He had multiple unsuccessful attempts to wean mechanical ventilation in the subsequent weeks. The phrenic nerve–conduction studies were then repeated 3 mos after injury, with improved CMAP amplitudes of 0.42 mV on the right side and 0.33 mV on the left side. After confirmation of improved inspiration by pulmonary function testing, weaning efforts were reintiated and he was successful in independent breathing during the next 4 days. His neurologic exam remained unchanged and consistent with C5, ASIA A, motor and sensory.
CASE 3

An 18-yr-old Caucasian male presented to us as an outpatient 11 mos after sustaining a cervical fracture dislocation during a motor vehicle accident. He was diagnosed with high cervical tetraplegia, C4 sensory level, ASIA A. He had been surgically stabilized and had undergone initial rehabilitation in another city. While at that institution, he had undergone multiple attempts at ventilator weaning without success. He had been advised that recovery of independent breathing was unlikely.

On our initial evaluation, he had a tracheostomy and portable positive-pressure ventilation incorporated into his wheelchair. His accompanying nurse reported that the patient did perform some voluntary respiratory movements while the tracheostomy tube was being changed. An electrodiagnostic study of the phrenic nerve and diaphragm was performed to further evaluate his respiratory status. Needle EMG of the diaphragm showed positive sharp waves and fibrillations bilaterally with decreased numbers of increased-duration polyphasic MUAPs firing at rates greater than 15 Hz during inspiratory recruitment. Phrenic nerve–conduction studies demonstrated that the CMAP amplitude on the right was severely diminished, but a normal response of 1.0 mV was obtained on the left side (Fig. 1). He was returned to the rehabilitation hospital for further assessment for weaning. This was accomplished successfully and he was decannulated. He did well, without further need for assisted mechanical ventilation.

DISCUSSION

In these three patients, we have documented improvements in CMAP amplitude in a phrenic motor nerve–conduction study that was associated with the functional recovery of breathing. This

<table>
<thead>
<tr>
<th>Time Since Injury</th>
<th>4 wks</th>
<th>2 mos</th>
<th>3 mos</th>
<th>4 mos</th>
<th>11 mos</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right</td>
<td>No response</td>
<td>8.9 msecs,</td>
<td>9.2 msecs,</td>
<td>7.9 msecs,</td>
<td>6.9 msecs,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.03 mV</td>
<td>0.42 mV*</td>
<td>0.49 mV*</td>
<td></td>
</tr>
<tr>
<td>Left</td>
<td>No response</td>
<td>8.2 msecs,</td>
<td>8.9 msecs,</td>
<td>8.4 msecs,</td>
<td>7.5 msecs,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.04 mV</td>
<td>0.33 mV*</td>
<td>0.13 mV*</td>
<td></td>
</tr>
<tr>
<td>Case 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right</td>
<td>9.6 msecs,</td>
<td>9.2 msecs,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.09 mV</td>
<td>0.42 mV*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left</td>
<td>9.3 msecs,</td>
<td>8.9 msecs,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.13 mV</td>
<td>0.33 mV*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right</td>
<td>No response*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left</td>
<td>8.8 msecs,</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.0 mV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Asterisks denote the amplitude values obtained from the diaphragm immediately before successful weaning from mechanical ventilation.

FIGURE 1 Bilateral diaphragm muscle CMAP recordings after ipsilateral phrenic nerve stimulation using surface electrodes in case 3. The left-side response (top) is at an amplifier gain of 0.5 mV per division and a sweep speed of 5 msecs per division. In the lower trace, there is no response seen on the right with the gain increased to 20 μV/cm. Soon after these recordings, the patient was successfully weaned from the ventilator.
correlates well with studies that have serially evaluated CMAP amplitude in other peripheral nerve lesions in association with functional recovery of the limb muscles. Late recovery of diaphragmatic function in high cervical SCI has also been reported. McKinley showed recovery of independent breathing with phrenic nerve CMAP amplitudes of 0.45 mV (left) and 0.21 mV (right) after 5 yrs of ventilator use. Four patients with initially small CMAP responses were seen to recover inspiratory ability during 6 mos or longer by Oo et al. Electrophysiologic recording of the improvement in diaphragmatic function after SCI has also been shown in animal models.

Phrenic nerve–conduction studies have also been used in humans as a method of selecting candidates for diaphragmatic pacing. Lieberman et al. have suggested serial testing for as long as 2 yrs after SCI because of variable results over time in his subjects. Overall, these results are similar to ours in that recovery occurred in some persons with initially small-amplitude CMAP responses. However, the findings are quite different in the situation of a desirable candidate for pacing who needs to have upper–motor neuron weakness (high SCI, above C3). These cases would demonstrate normal CMAP responses of the phrenic nerve in the absence of both fibrillations and voluntary MUAPs, in contrast to our cases, which showed a mixture of fibrillations and MUAPs.

The anterior horn cells for the motor neurons innervating the diaphragm reside at the third, fourth, and fifth cervical levels of the spinal cord. Therefore, a C4- or C5-level SCI can damage the anterior horn cells and nearby rootlets, which contribute axons to the phrenic nerves. This level of injury to the spine could also include extraspinal trauma to the phrenic nerve and its contributing components in addition to these intraspinal nerve injuries. Damage to the anterior horn cells and the phrenic nerve axons results in a decrease in the diaphragm CMAP and in fibrillations and MUAP abnormalities. Edema and anatomic distortions can also compress the phrenic nerves, possibly resulting in conduction block, which has the potential to resolve much more rapidly (often in a matter of weeks) than axonal regeneration, which can take many months.

The individuals presented here were studied to assess phrenic nerve function. Follow-up studies (in cases 1 and 2) were performed with the expectation of distal collateral sprouting of motor axons and possible early changes related to axonal regeneration. These expected methods of lower–motor neuron recovery, as well as the resolution of neuropraxic block, can lead to clinical improvement in diaphragm muscle strength and functional ventilation. Neurological and functional recovery can occur in SCI by other mechanisms, such as upper–motor neuron recovery. However, the individuals studied here displayed no other changes in motor or sensory neurologic status, and they remained classified as ASIA A (complete injury).

Phrenic nerve motor-conduction studies have been widely used in evaluation of respiratory function in SCI and other clinical conditions and can be performed safely and easily. The same evaluation serves both as an assessment of whether the patient is a candidate for use of a phrenic nerve or diaphragmatic stimulator for ventilation (e.g., the rare case of C3 or higher SCI) or whether there is lower–motor neuron injury and the possibility of peripheral nerve regeneration.

The cost and functional limitation of maintaining mechanical ventilation is an enormous burden. Aggressive weaning efforts are often made during the acute care stay in the early period after an SCI. Later, the possibility of neural recovery by either repair of conduction block or regeneration of motor axons must be considered as the natural course of the injury (as in case 3, and in McKinley’s). Case 3 demonstrates the value of a single study late after injury that can encourage a reattempt at ventilator weaning in a setting of the intact CMAP. In addition, the finding of active MUAPs by a needle EMG study of the diaphragm can demonstrate intact voluntary function of a portion of both the upper– and lower–motor neuron systems.

Electrodiagnostic testing does not replace pulmonary function testing, and further studies evaluating the correlation between phrenic CMAP amplitude and other traditional pulmonary measures, including fluoroscopic assessment of diaphragmatic motion (sniff test), could provide more enlightenment. Optimal care would include routine reevaluation of vital capacity or negative inspiratory force at the bedside, in addition to management of bronchial secretions and nutrition. Understanding the potential for peripheral reinnervation in injuries at the C4 and C5 levels can lead to appropriate anticipation of the improvement in motor nerve function and respiratory parameters. The recovery of muscle function here will often be reflected in the vital capacity and negative inspiratory force measures; however, these are subject to the patient’s anxiety and effort. EMG and nerve-conduction studies provide objective data but do not require the patient to experience the anxiety of going without ventilation. Correlation of vital capacity and negative inspiratory force results with electrodiagnosis can add confidence to the attempt to wean by helping the patient to understand that some of their nerve function has improved. The finding of early loss of the peripheral nerve–conduction response at the diaphragm does not ex-
clude a rare coexisting upper–motor neuron injury from which recovery is less likely, but the presence of combined injury can only be ascertained via serial study that includes the assessment of changes in the CMAPs and recorded MUAPs in addition to negative inspiratory force.

CONCLUSIONS

In these cases, the improvement of the phrenic nerve–response amplitudes were seen to correlate with clinical improvement and successful weaning from mechanical ventilation. All three individuals gained ventilator independence when at least one hemidiaphragm CMAP amplitude recovered to exceed 0.4 mV. Serial phrenic nerve–conduction studies can play a role in the evaluation and management of persons with lower–motor neuron dysfunction of the phrenic nerve in the situation of traumatic C4- or C5-level spine and SCI. It can also assist in understanding of the reinnervation process in these individuals. These cases are presented to demonstrate the potential benefit of phrenic nerve stimulation in assessing ventilator–dependent individuals with C4 or C5 SCI. Further studies are needed to correlate these findings with other commonly used parameters that predict successful ventilator weaning in these patients.

REFERENCES

Spinal Cord Infarction Secondary to Cocaine Use

ABSTRACT


A 27-yr-old woman recreationally inhaled cocaine. Several hours later, she noted chest tightness, back and neck pain, and later bilateral upper-extremity weakness. Physical examination revealed flaccid paresis of the upper extremities. Spasticity at 2 mos after injury, but no detectable weakness, developed in the lower extremities. Cocaine was detected in her urine. Magnetic resonance imaging showed hyperintensity in the anterior cervicothoracic spinal cord. Electrodiagnostic studies of the upper extremities were consistent with anterior horn cell death. Cocaine abuse is associated with cerebrovascular events; spinal cord effects are rarely reported. The patient seems to have an infarct in the anterior spinal artery distribution, with clinical, imaging, and electrodiagnostic findings of upper-extremity lower–motor neuron injury, accompanied by spasticity of the lower extremities. Gray matter has increased susceptibility to ischemia compared with white matter, producing flaccid weakness in the cervical region with isolated arm weakness. Although uncommon, cocaine abuse can cause spinal cord infarction.

Key Words: Case Report, Spinal Cord Infarction, Cocaine, Rehabilitation

The neurologic sequelae of cocaine use are typically intracranial. Some of the documented cases include intracerebral hemorrhage, subarachnoid hemorrhage, cerebral vasculitis, and seizures. We have found very few cases of cocaine abuse resulting in spinal cord infarction.

Mody et al. reported the case of a 24-yr-old man who smoked crack cocaine for the first time and suddenly developed a sharp pain between his shoulders that radiated up his neck. Within an hour, his hands felt numb and weak. He developed C5 tetraplegia with preservation of position and vibration sense within hours. Magnetic resonance imaging (MRI) done 6 wks later was normal. Sawaya and Kaminski reported the case of a 31-yr-old male who used intranasal cocaine chronically. He developed headache shortly after snorting two lines of cocaine. Shortly thereafter, he developed weakness in his upper extremities that progressed to tetraplegia and ventilator-dependent respiratory failure (VDRF) with preservation of position and vibration. Daras et al. reported the cases of three patients with cocaine use (nonspecific drug delivery) developing sudden paraplegia and anesthesia below the midthoracic level with preservation of posterior column function. Di Lazzaro et al. reported the case of 22-yr-old male...
who experienced cervical pain shortly after intranasal cocaine use; he later developed C2 level tetraplegia with VDRF. MRI revealed abnormal increase in signal from C2 to C7 on T2-weighted images, and follow-up MRI revealed abnormal intensities in the anterior horns of the gray matter. Di Lazzaro et al.5 also reported the case of a 37-yr-old male with weakness in the arm and legs immediately after injection with cocaine, with arm recovery the next day. One year later, his examination revealed hyperreflexia, spastic gait, and upgoing toes. MRI revealed hyperintensity in the anterior horns of the gray matter at C4 and C5.

Although there are only a few reported cases of spinal cord infarction, all cases seem to present as an infarction of the anterior spinal artery distribution with effect in the white and gray matter in the cervical region.

CASE DESCRIPTION

A 27-yr-old woman without past medical history recreationally snorted cocaine. Several hours later, she noted chest tightness, then back and neck pain, and, later, bilateral upper-extremity weakness.

When EMS arrived, she had vomited twice and developed a headache. On arrival to a local emergency department, her review of systems was negative for changes in vision, hearing, speech, or swallowing. Urine drug screen was positive for cocaine and marijuana. Her physical examination displayed marked weakness in her upper limbs with absent deep-tendon reflexes, and strength and reflexes were intact in the lower limbs. Full neurologic work-up, including MRI/angiogram of the brain, echocardiogram, lumbar puncture with cultures, and rheumatologic, infectious, and hypercoaguable serologies were negative. Electromyography/nerve conduction studies (EMG/NCS) were performed on the same day and were within normal limits. No infectious, inflammatory, embolic, or thrombotic cause was found. On the day of presentation, MRI of the spine revealed suspicion of hyperintense signal intensity at the cervicothoracic junction. Two days later, MRI of the spine revealed anterior spinal cord infarction without findings of occlusion, arterovenous malformation, or dissection on angiogram, felt to be secondary to anterior spinal artery vasospasm. T2-weighted MRI showed hyperintensity in the anterior cervicothoracic spinal cord with abnormal signal from C3 to T1 (Fig. 1).

On arrival to our acute rehabilitation unit, the patient had flaccid paresis of the upper limbs with preserved posterior column function, intact fine touch, and absent pin prick from C5 through L1. She also had a neurogenic bladder, which resolved, and very poor trunk stability. One month after injury, EMG/NCS was performed in the upper limbs. NCS revealed normal latencies and conduction velocities with diminished motor amplitudes. EMG revealed axonal denervating process with proximal reinnervation of the upper limbs, consistent with anterior horn cell death. At one month follow-up from rehabilitation, 2 mos from injury, she had spasticity, but no detectable weakness had developed in the lower limbs. She had full recovery in her proximal upper limbs but marked deficit in C7 to T1—in particular, bilaterally in the triceps and her left-hand intrinsics. At 8 mos from injury, she had consistent physical exam findings and was ambulating independently and working full time.

DISCUSSION

Cocaine abuse is well known to be associated with cerebrovascular events and stroke, perhaps secondary to vasospasm, hypertension, and cardiac problems. Spinal cord effects are less commonly reported. The patient described above seems to have sustained an infarct in the distribution of the anterior spinal artery, with clinical, imaging, and electrodiagnostic findings of upper-extremity lower-motor neuron injury, accompanied by spasticity of the lower extremities. The anterior spinal artery is the principal vascular supply to the spinal cord, originating in the most rostral portion from the vertebral arteries proximal to the union to form the

FIGURE 1 T2-weighted MRI with hyperintense signal intensity in the anterior spinal cord at levels C3 to T2, suggestive of ischemia, on the third day from onset of symptoms.
basilar artery. The anterior spinal artery receives input from six to nine radicular arteries. The major radicular artery is named the artery of Adamkiewicz. A lack of blood supply from the anterior spinal artery can cause an anterior cord syndrome, which affects both the white and gray matter, sparing the posterior funiculus and horns, which have blood supply from the posterior spinal arteries. The anterior horn may be preferentially affected because of greater susceptibility of gray matter to ischemia compared with white matter.6 Occurring in the cervical region, there is isolated arm weakness.7

CONCLUSION

Although there have been only a few reported cases of spinal cord infarction, the cases seem to present as infarction of the anterior spinal artery distribution.8 In this case, localized anterior horn cell death gave rise to flaccid weakness in the cervical region with isolated arm weakness. Although such effects are uncommon, cocaine abuse can cause spinal cord as well as cerebral infarction.

ACKNOWLEDGMENTS

This project was funded, in part, under grant number H133N000023 from the National Institute on Disability and Rehabilitation Research, Office of Special Education and Rehabilitative Services, United States Department of Education, Washington, DC.

REFERENCES

Belated Diagnosis of Congenital Myotonic Dystrophy in a Boy with Cerebral Palsy

ABSTRACT

We report a boy affected by both congenital myotonic dystrophy (CDM) and cerebral palsy (CP). He was born preterm and suffered from recurrent respiratory distress and respiratory tract infections. He was diagnosed with spastic diplegic CP. During the follow-up, long faces that lacked expression with a drooping mouth, of him and his mother, led to a detailed diagnostic evaluation. His electrodiagnostic testing demonstrated myotonic discharges. Molecular genetic tests revealed that the boy had 1240 CTG repeats and that his mother had 220 repeats, which confirmed the diagnosis of maternally transmitted CDM. Considering the increased obstetric risk of myotonic dystrophy (DM1), children of mothers with DM1 are susceptible to develop disorders such as CP. Clinical suspicion of comorbid CDM should be raised in children with CP who show features suggestive of DM1.

Key Words: Congenital Myotonic Dystrophy, Cerebral Palsy, CTG Repeats, Diagnosis

Myotonic dystrophy (DM1) is an autosomal-dominant neuromuscular disorder with multisystem involvement. The main clinical features are myotonia, muscle weakness preceding in distal muscles, cataracts, mental retardation, and cardiorespiratory abnormalities. The estimated incidence of DM1 is 1 in 8000 births, and its prevalence ranges from 2.1 to 14.3 per 100,000 worldwide. It is linked to a dynamic expansion of a CTG repeat in the 3’ untranslated region of a gene encoding DM1 protein kinase on chromosome 19q13.3.1 On the basis of age at onset, severity, and size of the CTG repeats, the phenotype is classified into four groups: mild, juvenile, classical, and congenital. Congenital myotonic dystrophy (CDM) is the most severe phenotypic expression and is observed in congenitally affected neonates born to classically affected mothers. The clinical picture of CDM includes respiratory distress, hypotonia, psychomotor retardation, and feeding difficulties.
Women with DM1 are known to be at increased risk for obstetric complications. Therefore, their offspring may be exposed to a higher chance of perinatal morbidity and mortality as well as disease inheritance.

In this study, we report a boy with both CDM and cerebral palsy (CP). He was diagnosed belatedly with CDM during regular follow-ups for CP. This is the first report of CDM associated with CP in a patient.

**CASE REPORT**

A currently 4-yr-old boy was the second born to nonconsanguineous Korean parents. He was delivered by a cesarean section because of preterm premature rupture of membrane at 33 wks of gestation. His body weight was 1940 g and his Apgar score was 5 at 1 and 5 mins. Shortly after birth, he showed progressive respiratory acidosis that required endotracheal intubation and ventilatory support until the eighteenth day. He was hospitalized three times in his first year for recurrent respiratory infections and respiratory distress. Laboratory investigations revealed no evidence of congenital infection or inborn errors of metabolism. Ultrasonography of the brain within the first month showed dilation of lateral ventricles, and brain magnetic resonance imaging at age 2 mos was suggestive of periventricular leukomalacia.

At age 5 mos, he was admitted to our hospital for further evaluation regarding his respiratory and developmental problems. Laryngological evaluation revealed no upper-airway problem. His motor development was clearly very delayed in that he still lacked head control. Neurological examination disclosed increased muscle tone in the lower extremities, which was not so apparent in the upper extremities. Increased deep-tendon reflex in the lower extremities and ankle clonus were observed bilaterally. Electroencephalogram results were normal, but follow-up brain magnetic resonance imaging showed atrophic change, delayed myelination, hypoplastic corpus callosum, and bilateral periventricular leukomalacia (Fig. 1).

Afterward, he was followed as an outpatient. He was diagnosed with spastic diplegic CP and referred to pediatric rehabilitation medicine outpatient clinic at 31 mos of age. The boy was first brought to our attention at age 40 mos because he displayed bilateral facial muscle weakness with a tent-shaped mouth (Fig. 2). His mother also

---

**FIGURE 1** FLAIR (A, B) and T2-weighted (C, D) axial images of brain magnetic resonance imaging at 5 mos of age. Patchy, high-signal lesions in bilateral periventricular white matter, lack of myelination, hypoplastic corpus callosum, and generalized brain atrophy are shown.
showed bilateral facial muscle weakness, nasal speech, and mild dysarthria. Given his and his mother's facial features and his delayed progress in therapy, an investigation of suspected DM1 was initiated. There was no remarkable family history of a neuromuscular disorder reported by his parents up to this point.

Cranial nerve examination revealed moderate facial weakness without ptosis. Grip myotonia and percussion myotonia were not observed in the hands. Manual muscle testing showed mild weakness of bilateral upper and lower distal limbs. Sensory examination was unremarkable. Deep-tendon reflexes in the lower limbs were hyperactive, and ankle clonus and the Babinski sign were observed bilaterally. He had planovalgus foot deformity and could walk unstably with ankle-foot orthoses. Video gait analysis revealed spastic walking on the toes with stiff knees and high-guard arm position.

On electrodiagnostic testing, motor- and sensory nerve-conduction studies were normal. Needle electromyography of limb musculature revealed typical waxing-waning “dive-bomber” myotonic discharges (Fig. 3). In several muscles, voluntary motor unit action potentials demonstrated reduced duration, reduced amplitude, and increased polyphasia.

Neurological examination of the patient’s mother revealed grip myotonia aggravated by cooling. In the obstetric history, she was gravida 2, para 2, and her first delivery had been performed by a cesarean section because of premature rupture of the membrane after 37 wks of gestation. Reduced fertility, polyhydramnios, or reduced fetal movement were not reported. His father and elder brother presented no signs of DM1.

The patient’s family underwent gene analysis for DM1. Analysis of the CTG expansion was performed on leukocyte DNA using the combination of Southern blot and polymerase chain reaction analysis. The boy had approximately 1240 CTG repeats, which was consistent with the clinical impression of CDM. The mother had 220 repeats, but his father and brother had fewer than 20.

A detailed family history was obtained (Fig. 4). The patient’s parents consented to scientific presentation of the genetic data and their pedigree. Ten members from three generations were clinically affected, and four of them were diagnosed with DM1. None of the patient’s paternal relatives were affected.

After establishing a diagnosis of CDM, related evaluations were performed and several consultations were obtained. Ophthalmologic examination was unremarkable. Echocardiography revealed a normal cardiac anatomy and ventricular function. Holter monitoring showed borderline PR interval prolongation. An oral glucose-tolerance test and creatine kinase were within normal limits. Speech-language evaluation revealed delayed language development, with the patient scoring below the 10th percentile for his age.

**DISCUSSION**

CP is a syndrome with movement and posture disorders caused by a nonprogressive injury to an
immature brain. The brain injury can occur at any point in the prenatal, perinatal, or postnatal periods. It is well known that the risk of obstetric complications is significantly increased in women with DM1. Obstetric risks associated with DM1 are preterm birth, perinatal mortality, polyhydramnios, operative delivery, cesarean section, placenta previa, and ectopic pregnancy.\textsuperscript{2,3} Like the mother in this report, most women with DM1 are not aware of their disease when they are pregnant,\textsuperscript{4} nor are they aware of the potential complications. Increased obstetric risks can lead their offspring to develop disorders such as CP. Neonates born to mothers with DM1 are also susceptible to being more severely affected because of the phenomenon called anticipation—\textsuperscript{5} that is, progressive severity and earlier manifestation of a disease in subsequent generations. Considering the obstetric risks and disease inheritance, the chances of CDM being associated with CP in a patient are significant.

Some clinical features are shared by both CDM and CP. This patient showed muscle weakness, delayed development, gastrointestinal problems, and oromotor dysfunction such as drooling, all of which are found in both disorders. Overlapping features are easily overlooked and can hinder diagnosis. Moreover, brain magnetic resonance imaging findings of DM1, such as white-matter abnormalities, ventricular dilation, and cerebral atrophy,\textsuperscript{6,7} are not disease specific, and they are similar to symptoms of CP. For these reasons, diagnosis of CDM in a patient with CP can be belated (as in our patient) or even missed.

To date, there has been no report of a patient with CP who was later diagnosed with CDM. In this study, it was the affected boy and his mother who brought the family to specific medical attention, and the diagnosis of the boy identified an extensive DM1 pedigree. What is imperative in the diagnosis is better awareness of the clinical features, and strong clinical suspicion. Needle electromyography, performed because of clinical suspicion, revealed myotonic potentials. Although clinical myotonia is not apparent in neonates and young children, myotonic discharges can sometimes be detected earlier on needle electromyography.\textsuperscript{8–10} Polyphasia also can be seen in addition to the usual myopathic electromyographic findings.\textsuperscript{8} This may actually help differentiate CDM from other neonatal myopathies.

Diagnosing additional neuromuscular problems in patients with CP is important for reducing disability and for formulating adequate therapeutic interventions such as medications for myotonia, more aggressive respiratory care programs, screening for associated conditions, and selection of the most appropriate goals for orthotic management and physical therapy, occupational therapy, and speech therapy. We provided our patient’s parents with information regarding prognosis, long-term...
therapeutic plans, and the need for evaluation of and genetic counseling for their relatives. This report emphasizes the importance of considering the diagnosis of DM1 in patients with CP who present with characteristically myopathic facial expressions and more profound psychomotor retardation than usually expected.

REFERENCES
Epidural Corticosteroid Injections
Precipitating Epidural Hematomas with Spinal Paresis
Myron M. LaBan, MD, MMSc, Gopi Kasturi, MD, and I-Ming Wang, MD

From the Departments of Physical Medicine and Rehabilitation (MML, GK) and Neuroradiology, William Beaumont Hospital, Royal Oak, Michigan (I-MW).

Epidural corticosteroid injections for intractable spinal pain, although often reputed to be innocuous, are not without risk. Both spinal epidural abscess formation and epidural hematomas\(^1\) have been reported after routine epidural corticosteroid injections for spinal radicular pain associated with disc herniations and/or spinal stenosis.

Four cases of epidural hematomas have been reported previously in the last 10 yrs; here, two additional cases are described.

**CASE 1**

A 36-yr-old female received three epidural steroid injections for cervical radiculopathy. After the second injection, she developed numbness and tingling in her hands. Magnetic resonance imaging demonstrated a hematoma from C3 to T3 (Fig. 1). An emergent surgical evacuation of the hematomas was completed the same day. She experienced both bladder and bowel incontinence. Strength was normal in the upper extremities, with minimal weakness in the lower extremities. Sensation was normal in the upper extremities but was diminished between T4 and T11. At discharge, she was able to ambulate independently for 150 feet with a rolling walker.

**CASE 2**

A 79-yr-old female was admitted to the hospital for an elective epidural steroid injection with symptoms of intratable lumbar spinal stenosis. Coumadin, which had been prescribed for atrial fibrillation, was discontinued, and she was started on IV heparin. Two days later, she was given an epidural corticosteroid injection. She subsequently developed paraparesis. A computed tomographic myelogram demonstrated a hematoma from T2 to T8 with cord compression (Fig. 2). Her anticoagulation was reversed and a decompressive thoracic laminectomy was performed the next day. The patient remained paraparetic with both a neurogenic bowel and bladder. She was discharged to a subacute rehabilitation facility after 1 mo of acute rehabilitation, able to transfer and propel her wheelchair 150 feet.

Reports of epidural hematomas after epidural corticosteroid injections are relatively infrequent, asymptomatic, and/or unreported.

Since Jackson\(^2\) first described the epidural hematoma in 1869, numerous etiologic agents have been reported, including surgery, trauma, anticoagulation, arterial venous malfor-
mations, pregnancy, anticoagulation therapy, and lumbar puncture. Disruption of the integrity of the internal vertebral plexus of Batson has been cited in these instances as a potential source of bleeding.\textsuperscript{3} As demonstrated by Tarlov et al.,\textsuperscript{4} recovery of spinal neurological function is dependent on both the magnitude and duration of the cord compression antecedent to surgery. Anticipating the risk of an epidural bleed is contingent on identifying and managing predisposing risk factors including, among others, anticoagulation therapy, and also on weighing the benefits of the procedure. Preprocedure disclosure should continue to acknowledge the relatively rare but potentially devastating consequences of either an epidural spinal hematoma and/or infection.

REFERENCES

4. Tarlov IM, Klinger H, Vitale S: Spinal cord compression studies: 1: experimental techniques to produce acute and gradual compression. \textit{Arch Neurol Psychiatry} 1953;70:813–19

\textbf{FIGURE 2} Lumbar spine epidural hematoma (marked by asterisk).