## Guest Editorial

<table>
<thead>
<tr>
<th></th>
<th>Editorial Board</th>
<th>Page CO2</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Table of Contents</td>
<td>Page A1</td>
</tr>
<tr>
<td>3</td>
<td>Information for Readers</td>
<td>Page A2</td>
</tr>
</tbody>
</table>

### When Evidence and Practice Collide

<table>
<thead>
<tr>
<th></th>
<th>Original Articles</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>When Evidence and Practice Collide</td>
</tr>
<tr>
<td></td>
<td>Pages 551-553</td>
</tr>
<tr>
<td></td>
<td>Robert D. Mootz</td>
</tr>
<tr>
<td>5</td>
<td>JMPT Highlights</td>
</tr>
<tr>
<td></td>
<td>Page 554</td>
</tr>
</tbody>
</table>

### Cost-Effectiveness of Medical and Chiropractic Care for Acute and Chronic Low Back Pain

<table>
<thead>
<tr>
<th></th>
<th>Original Articles</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>Cost-Effectiveness of Medical and Chiropractic Care for Acute and Chronic Low Back Pain</td>
</tr>
<tr>
<td></td>
<td>Pages 555-563</td>
</tr>
<tr>
<td></td>
<td>Mitchell Haas, Rajiv Sharma and Miron Stano</td>
</tr>
</tbody>
</table>

### Effects of a Managed Chiropractic Benefit on the Use of Specific Diagnostic and Therapeutic Procedures in the Treatment of Low Back and Neck Pain

<table>
<thead>
<tr>
<th></th>
<th>Original Articles</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>Effects of a Managed Chiropractic Benefit on the Use of Specific Diagnostic and Therapeutic Procedures in the Treatment of Low Back and Neck Pain</td>
</tr>
<tr>
<td></td>
<td>Pages 564-569</td>
</tr>
<tr>
<td></td>
<td>Craig F. Nelson, R. Douglas Metz and Thomas LaBrot</td>
</tr>
</tbody>
</table>

### A Cross-Sectional Study Comparing Pain and Disability Levels in Patients With Low Back Pain With and Without Transitional Lumbosacral Vertebrae

<table>
<thead>
<tr>
<th></th>
<th>Original Articles</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>A Cross-Sectional Study Comparing Pain and Disability Levels in Patients With Low Back Pain With and Without Transitional Lumbosacral Vertebrae</td>
</tr>
<tr>
<td></td>
<td>Pages 570-574</td>
</tr>
<tr>
<td></td>
<td>Cynthia K. Peterson, Jennifer Bolton, William Hsu and Angela Wood</td>
</tr>
</tbody>
</table>

### Erector Spinae and Quadratus Lumborum Muscle Endurance Tests and Supine Leg-Length Alignment Asymmetry: An Observational Study

<table>
<thead>
<tr>
<th></th>
<th>Original Articles</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>Erector Spinae and Quadratus Lumborum Muscle Endurance Tests and Supine Leg-Length Alignment Asymmetry: An Observational Study</td>
</tr>
<tr>
<td></td>
<td>Pages 575-581</td>
</tr>
<tr>
<td></td>
<td>Gary A. Knutson and Edward Owens</td>
</tr>
</tbody>
</table>
| 10. | Effect of High-Intensity Strength-Training on Functional Measures of Balance Ability in Balance-Impaired Older Adults  
*Pages 582-590*  
Jennifer A. Hess and Marjorie Woollacott |
|---|---|
| 11. | Measurement of the Pressure Applied During Motion Palpation and Reliability for Cervical Spine Rotation  
*Pages 591-596*  
Justin Marcotte, Martin C. Normand and Pierre Black |
*Pages 597-603*  
Deed E. Harrison, Jason W. Haas, Rene Cailliet, Donald D. Harrison, Burt Holland and Tadeusz J. Janik |
| 13. | Validation of the Spin-T Goniometer, a Cervical Range of Motion Device  
*Pages 604-609*  
Shabnam Agarwal, Garry T. Allison and Kevin P. Singer |
| 14. | Validity of the Lateral Gliding Test as Tool for the Diagnosis of Intervertebral Joint Dysfunction in the Lower Cervical Spine  
*Pages 610-616*  
Cesar Fernández-de-las-Peñas, Cristobal Downey and Juan Carlos Miangolarra-Page |

**Literature Reviews**

*Pages 617-622*  
Michael T. Haneline and Gary N. Lewkovich |
| 16. | The Biomechanical and Clinical Significance of the Lumbar Erector Spinae Flexion-Relaxation Phenomenon: A Review of Literature  
*Pages 623-631*  
Christopher J. Colloca and Richard N. Hinrichs |
| 17. | Labral Injuries of the Hip: A Review of Diagnosis and Management  
*Pages 632.e1-632.e8*  
Matt Schmerl, Henry Pollard and Wayne Hoskins |

**Case Report**

| 18. | Pathological Cervical Fracture after Spinal Manipulation in a Pregnant Patient  
*Pages 633-636*  
Alfred Schmitz, Goetz Lutterbey, Lars von Engelhardt, Marcus von Falkenhausen and Michael Stoffel |
Letters to the Editor

19. The Bournemouth Questionnaire: Can it be Used to Monitor and Predict Treatment Outcome in Chiropractic Patients With Persistent Low Back Pain?
   Page 637
   Jennifer E. Bolton

20. Response to Letter to the Editor by Bolton
   Pages 637-638
   Kristian Larsen and Charlotte Leboeuf-Yde

   Pages 638-639
   Ronald J. Feise
# TABLE OF CONTENTS

## GUEST EDITORIAL
551 When Evidence and Practice Collide  
*Robert D. Mootz, DC*

## ORIGINAL ARTICLES
555 Cost-Effectiveness of Medical and Chiropractic Care for Acute and Chronic Low Back Pain  
*Mitchell Haas, DC, MA, Rajiv Sharma, PhD, and Miron Stano, PhD*

564 Effects of a Managed Chiropractic Benefit on the Use of Specific Diagnostic and Therapeutic Procedures in the Treatment of Low Back and Neck Pain  
*Craig F. Nelson, DC, MS, R. Douglas Metz, DC, and Thomas LaBrot, DC*

570 A Cross-Sectional Study Comparing Pain and Disability Levels in Patients With Low Back Pain With and Without Transitional Lumbosacral Vertebrae  
*Cynthia K. Peterson, DC, MMEd, Jennifer Bolton, PhD, MSc, William Hsu, DC, and Angela Wood, DCR(R)*

575 Erector Spinae and Quadratus Lumborum Muscle Endurance Tests and Supine Leg-Length Alignment Asymmetry: An Observational Study  
*Gary A. Knutson, DC, and Edward Owens, DC*

582 Effect of High-Intensity Strength-Training on Functional Measures of Balance Ability in Balance-Impaired Older Adults  
*Jennifer A. Hess, DC, MPH, PhD, and Marjorie Woollacott, PhD*

591 Measurement of the Pressure Applied During Motion Palpation and Reliability for Cervical Spine Rotation  
*Justin Marcotte, DC, Martin C. Normand, DC, PhD, and Pierre Black, MSc*

597 Concurrent Validity of Flexicurve Instrument Measurements: Sagittal Skin Contour of the Cervical Spine Compared With Lateral Cervical Radiographic Measurements  
*Deed E. Harrison, DC, Jason W. Haas, DC, Rene Cailliet, MD, Donald D. Harrison, PhD, DC, MSc, Burt Holland, PhD, and Tadeusz J. Janik, PhD*

604 Validation of the Spin-T Goniometer, a Cervical Range of Motion Device  
*Shabnam Agarwal, MSc, Garry T. Allison, PhD, PT, and Kevin P. Singer, PhD, PT*

610 Validity of the Lateral Gliding Test as Tool for the Diagnosis of Intervertebral Joint Dysfunction in the Lower Cervical Spine  
*Cesar Fernández-de-las-Peñas, PT, Cristobal Downey, PT, MSc, and Juan Carlos Miangolarra-Page, MD, PhD*

## LITERATURE REVIEWS
617 An Analysis of the Etiology of Cervical Artery Dissections: 1994 to 2003  
*Michael T. Haneline, DC, MPH, and Gary N. Lewkovich, DC*

623 The Biomechanical and Clinical Significance of the Lumbar Erector Spinae Flexion-Relaxation Phenomenon: A Review of Literature  
*Christopher J. Colloca, DC, and Richard N. Hinrichs, PhD*

632 Labral Injuries of the Hip: A Review of Diagnosis and Management  
*Matt Schmerl, MChiro, Henry Pollard, Grad DC, MSPORTSc, PhD, and Wayne Hoskins, MChiro*

## CASE REPORT
633 Pathological Cervical Fracture After Spinal Manipulation in a Pregnant Patient  
*Alfred Schmitz, MD, Goetz Lutterbey, MD, Lars von Engelhardt, MD, Marcus von Falkenhausen, MD, and Michael Staffel, MD*

## LETTERS TO THE EDITOR
637 The Bournemouth Questionnaire: Can it Be Used to Monitor and Predict Treatment Outcome in Chiropractic Patients With Persistent Low Back Pain?  
*Jennifer E. Bolton, PhD, MA Ed, ILTM, FCC(Hon)*

637 Response to Letter to the Editor by Bolton  
*Kristian Larsen, PT, MPH, and Charlotte Leboeuf-Yde, DC, MPH, PhD*

638 False Negative Magnetic Resonance Imaging Results: A Report of 2 Cases  
*Ronald J. Feise, DC*

---

*Full-text online access to *JMPT* is available for all print subscribers. See page 569 for details.*
Questions or Comments?
Correspondence regarding editorial matters should be addressed to the Editor: Claire Johnson, DC, MSEd, Journal of Manipulative and Physiological Therapeutics, 200 E. Roosevelt Rd, Lombard, IL 60148-4583.

Instructions for Authors
At any time the instructions for authors can be viewed at www.mosby.com/jmpt or can be obtained from the editor. Authors should consult these instructions before submitting manuscripts.

Advertising Communications
Contact Periodical Advertising Sales Services at 314-453-4307 or 1-800-325-4177, ext 4307.

Business Communications

Subscriptions
Subscriptions may begin at any time. To enter a subscription to the Journal of Manipulative and Physiological Therapeutics, call 800-654-2452 or 407-345-4000; fax 407-363-9661; or e-mail elspcs@elsevier.com. Remittances made by check, draft, post office, or express money order should be in US funds, drawn through a US bank, made payable to this Journal, and sent to Elsevier Inc, Subscription Customer Service, 6277 Sea Harbor Dr, Orlando, FL 32887, USA.

2005 US subscription rates: individual, $119.00; institution, $212.00; student and resident, $59.00; single issue, $28.00. Outside of the US and possessions: individual, $157.00; institution, $253.00; student and resident, $78.00. Canadian customers, please add 7% GST to international prices. Prices subject to change without notice. Subscription rates include supplements and full-text online access. To receive student/resident rate, orders must be accompanied by name of affiliated institution, date of term, and the signature of program/residency coordinator on institution letterhead. Orders will be billed at individual rate until proof of status is received.

Single-copy prices will be charged on missing issues older than 3 months (6 months international) from mail date. Back issues generally are available for the previous 5 years.

Indexing/Abstracts
The Journal is indexed or abstracted in Index Medicus, Excerpta Medica, Index to Chiropractic Literature, Current Contents/Clinical Medicine, Russian Academy of Sciences, and BS/CCML Database. Volume index appears in the December issue.

Reprints
To order author reprints, contact Periodical Reprints at 800-325-4177, ext 4350, or 314-453-4350; fax 314-579-3358; e-mail: author.reprints@elsevier.com. To order 100 or more reprints for educational, commercial, or promotional use, contact the Commercial Reprints Department, Elsevier Inc, 360 Park Ave South, New York, NY 10010-1710. Fax: 212-462-1935; e-mail: reprints@elsevier.com. Reprints of single articles available online may be obtained by purchasing Pay-Per-View access for $30.00 per article on the journal Web site, www.mosby.com/jmpt.

Microform
Microform edition of the Journal is available from ProQuest Information and Learning, 300 N Zeeb Rd, Ann Arbor, MI 48106-1346.

Copyright
© 2005 National University of Health Sciences. All rights reserved.

This Journal and the individual contributions contained in it are protected under copyright by National University of Health Sciences, and the following terms and conditions apply to their use:

Photocopying
Single photocopies of single articles may be made for personal use as allowed by national copyright laws. Permission of the Publisher and payment of a fee is required for all other photocopying, including multiple or systematic copying, copying for advertising or promotional purposes, resale, and all forms of document delivery. Special rates are available for educational institutions that wish to make photocopies for nonprofit educational classroom use.

Permissions may be sought directly from Elsevier’s Rights Department in Philadelphia, PA, USA: phone (+1) 215 239 3804, fax: (+1) 215 239 3805, e-mail healthpermissions@elsevier.com. Requests may also be completed online via the Elsevier homepage (http://www.elsevier.com/locate/permissions).

In the USA, users may clear permissions and make payments through the Copyright Clearance Center, Inc., 22 Rosewood Drive, Danvers, MA 01923, USA: phone (+1) (978) 7508400, fax: (+1) (978) 7504744, and in the UK through the Copyright Licensing Agency Rapid Clearance Service (CLARCS), 9 Tottenham Court Road, London W1P 0LP, UK; phone: (+44) 20 7631 5555; fax: (+44) 20 7631 5500. Other countries may have a local reprographic rights agency for payments.

Derivative Works
Subscribers may reproduce tables of contents or prepare lists of articles including abstracts for internal circulations within their institutions. Permission of the Publisher is required for resale or distribution outside the institutions.

Permission of the Publisher is required for all other derivative works, including compilations and translations.

Electronic Storage or Usage
Permission of the Publisher is required to store or use electronically any material contained in this journal, including any article or part of an article.

Except as outlined above, no part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without prior written permission of the Publisher.

Address permissions requests to: Elsevier Rights Department, at the fax and e-mail addresses noted above.
"Until now, we believed that the best way to transmit knowledge from its source to its use in patient care was to first load the knowledge into human minds... and then expect those minds, at great expense, to apply the knowledge to those who need it. However, there are enormous ‘voltage drops’ along this transmission line for medical knowledge."

-Lawrence Weed, 1999

It sounds good, doesn’t it? Evidence-based practice (EBP)—Sackett has characterized it as the conscientious and judicious use of current best evidence in making decisions about the care of individual patients. Key to his conceptions are that individual clinical expertise, best research evidence, and patient values are all factored into clinical decision making. Bolton has furthered the ideal with her contentions that the evolution of EBP requires maturation of both research and practice methodologies to better address the issue of evidence suitability. Key to her conceptions are that EBP specifically involves matching knowledge gleaned from scientifically derived information to individual patient settings, that is, encouraging individual physicians to explicitly seek out and factor in scientifically gained knowledge into daily clinical decision making. Of course, this entails overcoming the inertia of practicing within the “comfort zones” that we have developed based on our training.

Although resources now abound to assist clinicians in the process of incorporating evidence into practice, the reality of information overload still prevails. As of 2004, the National Library of Medicine added almost 11,000 new articles per week to its databases. Just to stay current in internal medicine, an internist would need to read some 20 articles per day 365 days a year. When I began my career, I can proudly proclaim that I owned every chiropractic textbook published by a mainstream publisher and knew personally every chiropractic researcher in the indexed literature. Fortunately, for the science of chiropractic, that no longer holds true. Thus, “relevance retrievability” from the mass of literature has become acutely critical.

Furthermore, the user-friendly secondary information sources such as evidence summaries and clinical practice guidelines we rely on to offset information overload suffer from every limitation human nature conjures up including error, aging, bias, and misinterpretation. Superimposed on all of the impediments individual physicians must overcome are administrative hassles that spring forth from the business and regulatory constraints of contemporary health care. Dichotomous decisions have to be made about what to pay for with pooled public or private resources, balanced among competing demands of the market place. Fig 1 summarizes evidence ideals and realities faced by both clinicians and policy makers. In fact, it is this policy/practice interface where evidence and practice collide. The following 3 conceptual strategies illustrate our challenges.

“EVIDENCE APPROPRIATISM”

This is how EBP was meant to be. Evidence is used by all to better inform the choices patients have in making clinical decisions with their physicians, not to regulate them. End users have an appreciation of the strengths and limitations of varying qualities of evidence. Evidence suitability trumps evidence hierarchies, and the end users have realtime access to high-quality information. Guidelines are living documents refined regularly with new information and experience. Evidence is a tool to improve practice specifically and how we administer health care generally. It also is a tool that may require you to change what you do.

“EVIDENCE NIHILISM”

This represents the perspective that one cannot act until definitive evidence is available. From such a vantage point, the absence of evidence qualifies as evidence against, as does conflicting evidence. Although it sounds extreme, there are situations (such as when making a coverage decision regarding a procedure with high risk of adverse outcome) where more rigorous standards may be appropriate. However, this approach can sometimes be applied in default fashion to the detriment of various clinical situations.

“EVIDENCE AGENDAISM”

This strategy reflects the selective use of evidence to bolster one’s preconceived notions. For example, a payer might seek out and act on evidence that favors their business need while ignoring that which does not. Capping
Evidence ideals in clinical practice:
- Evidence summaries provide rapid access to critically appraised information of high quality
- High-quality evidence helps to inform choices that doctors and patients must make about care options
- An optimal balance of training, experience, and current research optimizes patient outcomes and value

Evidence ideals in policy:
- Evidence helps policy makers and regulators assure that high quality, effective care is provided and paid for with community resources
- Evidence helps avoid wasting resources on low quality or harmful care

Evidence realities in clinical practice:
- Evidence is often limited in quality, quantity, applicability to populations treated
- Trade associations frequently select evidence that underscores their viewpoint and ignores conflicting evidence
- Experience and judgment must be applied when relevant evidence is absent

Evidence realities in policy:
- Payers often use evidence-summaries or interpretations of evidence to justify dichotomous adjudicative stop-care points and cost-containment strategies
- Burden of justification of care decision in absence of evidence is (often unfairly) placed on the individual provider

chiropractic benefit or restricting coverage for known best practices under the auspices of “following a guideline” comes to mind. It may be a practical business decision to make, but “blaming the evidence” illustrates this strategy. Likewise, providers may promulgate and promote studies that emphasize a miniscule, marginal advantage (eg, a small improvement in pain in the absence of any functional improvement) or benefits out of context. Citing beneficial cost studies that used assumptions of noncoverage in a Canadian province to a US payer that already has a moderately robust chiropractic benefit might be an example. “Agendaism” may be naïve or overt. Of course, when evidence does appropriately challenge one’s own preconceived notions, the challenged party might wrongly assume agendaism on the part of the challenger.

The evidentiary contest in which the health care system now finds itself requires all parties to improve their understanding and application of evidence. Adapting behavior to aptly grapple with what evidence tells us is perhaps the highest priority. Clinicians must build time and routine into their workday to consider evidence in individual care decisions, as well as constructively engage administrative efforts to apply evidence to their decision making. Because it is change, it is challenging, but the evolution is straightforward.

Within a clinical setting, becoming a consumer of research information is really as easy as searching a free online database such as PubMed for articles on a couple clinical conditions seen in the past week. Reading through abstracts on recent literature may reveal new best practices or shed light on practices that are not useful. Development of an “evidence culture” will prepare chiropractors for adapting to the changes facing health care. Fig 2 outlines several examples of what practitioners can do, individually and collectively, to increase their comfort level with evidence and, thus, recognize its value and limitations for practice.

1. Develop an evidence comfort level
   - Create monthly journal clubs & in-service training with peers and colleagues
   - Seek out information summaries. For example:
     o Systematic Reviews and Meta-analyses
     o Cochrane Library, Bandolier
     o Clinical Guidelines
     o Narrative Reviews
   - Search for “evidence-based case reports”
   - Search for “Best Evidence Topics (BETs)” and “Critically Appraised Topics (CATs)”
   - Read and subscribe to literature in your field

2. Begin patient-centered performance tracking
   - Pick conditions common to your practice and chart how well you do with your patients over a few months. For example:
     o Measure the speed to functional recovery (return to work, activity outcome, functional outcome, etc)
     o Review compliance with “best practices” reported in the literature

3. Initiate EBP focused partnership opportunities
   - Initiate intra-professional relations with your county or state society groups. For example:
     o Pick a clinical topic and rotate among colleagues presenting what research is effective for various problems (eg, taping for knee injuries, safe return to work, etc.)
   - Initiate inter-professional relations with other local health providers. For example:
     o Sponsor an exercise and rehabilitation grand rounds with local physical therapy groups
     o Establish patient selection criteria for advanced imaging with a local radiology group
   - Develop best-practice workshops in provider-payer forums
     o Work with local or state societies to sponsor workshops with medical directors or payers on how to better document and communicate care progress for fair adjudication

Fig 1. Ideals and realities about EBP.

Fig 2. Strategies for practitioners to evolve in the EBP era.
To operationalize EBP, it is critical to focus on “patient-oriented evidence that matters.” That simply means that the outcomes one aims to improve are truly relevant to the patient. For example, a clinical finding such as segmental range of motion that is believed by the physician to be useful may be of far less interest to the patient or society at large than a functional performance outcome such as how long they can perform work with out pain. Physiologic measures may be of interest, but the value and tangible benefit to the consumer-patient as well as the customer-payer must be clear.

Evidence-based practice can also help to refine care processes to reduce mistakes. From a quality-improvement perspective, reduction and elimination in defects (misuse, underuse, and overuse) to diminish waste and inefficiency will not only benefit patients, but also will enhance the “competitiveness” of the practitioner in the community. For examples, see the Bridges to Excellence Web site (www.bridgestoexcellence.org/bte/index.html). There is increasing availability of comparative provider performance data now being made available to consumers. This trend will likely increase and is envisioned as a way to contribute to increased accountability as well as quality improvement.

Policy makers must confront the dilemmas that evidence can only inform the adjudication and policy decisions they must undertake. Evidence cannot make decisions for them. Researchers, more than ever before, must recognize that subtleties such as which outcome will become a primary hypothesis or which exclusion criteria are adopted may inadvertently drive a large-scale policy decision that impacts broader coverage.

Evidence-based practice is a contemporary reality impacting clinician, scientist, and policy maker alike. Inherent in all that EBP stands for is bettering patient outcomes and reducing expenditures on ineffective care. Evidence-based practice is still maturing, and infrastructures for supporting and adapting to it are still evolving. With all its inherent challenges, EBP remains the best hope in overcoming Weed’s “voltage drops.”

ACKNOWLEDGEMENT

Presentations by Jennifer Bolton, PhD, of Anglo-European College of Chiropractic and Douglas Weeks, PhD, of St Luke’s Rehabilitation Hospital have provided inspiration for much of this discourse. However, any and all limitations, errors, and oversight in this commentary are exclusively my own.

REFERENCES

Cervical spine fracture.

Schmitz et al (p. 633) report on an unusual case of immediate onset of neck pain after cervical manipulation and suggest clinical considerations for similar cases.

Do transitional spinal segments cause pain?

Peterson et al (p. 570) investigate if male and female patients with and without transitional lumbosacral vertebrae present with differences in pain and disability levels.

Acetabular labral tears.

Schmerl et al (p. 632) review the literature and report on the current knowledge of the diagnosis and treatment of acetabular labral tears.

Measuring cervical lordosis.

Harrison et al (p. 597) compare flexicurve surface contour measurements of the cervical spine with radiographic measurements of cervical lordosis.

The force of palpation.

Marcotte et al (p. 591) measure the pressure applied during motion palpation for cervical spine rotation and look at the effect of the pressure of palpation on the interexaminer reliability.

Helping geriatric patients.

Hess and Woollacott (p. 582) evaluate the effect of a 10-week high-intensity strength-training program targeting lower extremity muscles for the purpose of improving postural control in balance-impaired older adults.

A closer look at leg length.

Knutson and Owens (p. 575) investigate the association between a commonly used sign of joint dysfunction, supine leg-length alignment asymmetry, and the endurance of the erector spinae and quadratus lumborum muscles.

Is palpation valid?

Fernández-de-las-Peñas et al (p. 610) compare the lateral gliding test for intervertebral joint dysfunctions of the lower cervical spine to radiological assessment for patients with mechanical neck pain.

Searching for the cause of cervical artery dissections.

Haneline and Lewkovich (p. 617) provide an estimate of the etiologic breakdown of cervical artery dissections reported in the previous decade.

Chiropractic in managed care.

Nelson et al (p. 564) investigate if the inclusion of a chiropractic benefit for the treatment of low back and neck pain resulted in a reduction in the rates of surgery, advanced imaging, inpatient care, and plain-film radiographs.

Is high tech better?

Agarwal et al (p. 604) compare measurements of the simple, clinical cervical spine Spin-T goniometer with that of a high-resolution motion-tracking system for cervical ranges of motion.

Costs for the care of low back pain.

Haas et al (p. 555) find that chiropractic care is relatively cost-effective for the treatment of chronic low back pain, whereas chiropractic and medical care are comparable for patients with the acute condition.

Flexion-relaxation phenomenon.

Colloca and Hinrichs (p. 623) review the biomedical literature to ascertain the biomechanical and clinical significance of the lumbar erector spinae flexion-relaxation phenomenon.
OBJECTIVE ARTICLES

COST-EFFECTIVENESS OF MEDICAL AND CHIROPRACTIC CARE FOR ACUTE AND CHRONIC LOW BACK PAIN

Mitchell Haas, DC, MA, a Rajiv Sharma, PhD, b and Miron Stano, PhD c

ABSTRACT

Objectives: To identify relative provider costs, clinical outcomes, and patient satisfaction for the treatment of low back pain (LBP).

Methods: This was a practice-based, nonrandomized, comparative study of patients self-referring to 60 doctors of chiropractic and 111 medical doctors in 51 chiropractic and 14 general practice community clinics over a 2-year period. Patients were included if they were at least 18 years old, ambulatory, and had low back pain of mechanical origin (n = 2780). Outcomes were (standardized) office costs, office costs plus referral costs for office-based care and advanced imaging, pain, functional disability, patient satisfaction, physical health, and mental health evaluated at 3 and 12 months after the start of care. Multiple regression analysis was used to correct for baseline differences between provider types.

Results: Chiropractic office costs were higher for both acute and chronic patients (P < .01). When referrals were included, there were no significant differences in either group between provider types (P > .20). Acute and chronic chiropractic patients experienced better outcomes in pain, functional disability, and patient satisfaction (P < .01); clinically important differences in pain and disability improvement were found for chronic patients only.

Conclusions: Chiropractic care appeared relatively cost-effective for the treatment of chronic LBP. Chiropractic and medical care performed comparably for acute patients. Practice-based clinical outcomes were consistent with systematic reviews of spinal manipulation efficacy: manipulation-based therapy is at least as good as and, in some cases, better than other therapies. This evidence can guide physicians, payers, and policy makers in evaluating chiropractic as a treatment option for low back pain. (J Manipulative Physiol Ther 2005;28:555-563)

Key Indexing Terms: Low Back Pain; Chiropractic Care; Medical Care; Cost-Effectiveness

Healthcare costs for the treatment of back pain are substantial. A recent incremental spending model for the United States indicates that the additional costs associated with back pain patients represent 2.5% of national health care expenditures (a value expected to reach to $48 billion for 2005). With wide variations in spending patterns across patients with different clinical, socioeconomic, and demographic characteristics, the authors concluded that more cost-effective and targeted treatments could produce significant health care savings. With most adults experiencing back pain at some point in their lives, such treatments would represent an important public health improvement.

Because nonmedical providers, most notably chiropractors, provide a substantial portion of care for patients with low back pain (LBP), the relative efficacy and cost-effectiveness of chiropractic and medical care have emerged as important issues in the broader debate on evidence-based medicine. The growth of managed care and other gatekeeper mechanisms that restrict patient access to both medical

a Professor and Dean of Research, Center for Outcomes Studies, Western States Chiropractic College, Portland, Ore.
b Assistant Professor, Department of Economics, Portland State University, Portland, Ore.
c Professor, Department of Economics, School of Business Administration, Oakland University, Rochester, Mich.

Sources of support: This study was supported by the Health Resources and Services Administration, Department of Health and Human Services, Rockland, Md (grant no. R18 AH10002) and by a competitive challenge grant from the Foundation for Chiropractic Education and Research, Norwalk, Iowa (grant no. 940502) with funds donated to the foundation by the National Chiropractic Mutual Insurance Corporation, Des Moines, Iowa.

Submit requests for reprints to: Mitchell Haas, DC, MA, Center for Outcomes Studies, Western States Chiropractic College, 2900 NE 132nd Avenue, Portland, OR 97230 (e-mail: mhaas@wschiro.edu).


Copyright © 2005 by National University of Health Sciences. doi:10.1016/j.jmpt.2005.08.006
specialists and nonmedical providers have heightened the need for additional evidence that could be used to better allocate health care dollars.

There is a considerable body of randomized trials on the efficacy of spinal manipulation for the treatment of LBP; this evidence is summarized in the most recent systematic reviews. Assendelft et al concluded that manipulation is superior to placebo and sham procedures but no better than other commonly used therapies. In a companion review, Cherkin et al concluded that manipulation is at least as effective as other therapies. Bronfort et al found no treatment superior to manipulation and concluded that manipulation is a viable treatment option for acute and chronic LBP. More recent trials have also supported efficacy of spinal manipulation. Our practice-based, nonrandomized comparative study showed a clinically important advantage for chiropractic care over medical care for chronic patients and a marginal advantage for acute patients.

Early cost studies showed both lower and higher costs for chiropractic care than for other interventions. These studies had diverse designs, payment types, and analytic methods. In a comprehensive literature review of occupational LBP, Baldwin et al concluded that chiropractic and medical care are equally effective, but because of conflicting evidence and methodologic shortcomings, evidence for relative cost-effectiveness is inconclusive. No studies combined sufficient sample size, confounder controls, and high-quality cost data. Solomon et al were similarly critical of study methodology.

Since these reviews, a large managed care network in California found that members who received chiropractic coverage had 12% lower annual health care expenditures (1.6% lower after adjusting for member risk characteristics) than members without the coverage. Patients with the chiropractic benefit had lower back pain cost per episode of back pain, as well as lower rates of surgery and hospitalization. A randomized trial in the United Kingdom found that spinal manipulation alone or with exercise can be the best strategy, so long as a quality-adjusted life-year is valued above £3800 (then approximately US $5700). Another randomized trial in Sweden reported that costs and outcomes were generally similar for physiotherapy and chiropractic. The authors concluded that the therapies were equivalent from a cost-effectiveness perspective.

A preliminary report from our study indicated that mean direct in-office costs of patients treated by chiropractors were 74% higher (median, 39% higher) than those treated by medical physicians. However, the report did not distinguish acute from chronic patients, and cost and outcomes comparisons were unadjusted for baseline group differences. A potentially more important limitation was the exclusion of referral and advanced imaging costs.

This report fills these gaps by applying multiple regression analysis to cost as well as outcomes data. It contrasts analysis of office costs with and without costs of referral and advance imaging. Analysis was conducted separately for acute and chronic patients with LBP, in accordance with the original study design. It also includes a more extensive set of patient outcomes measures that permit estimation of incremental cost-effectiveness ratios. Analysis was conducted for one short-term and one long-term time point; 3 and 12 months were chosen a priori for this report.

Methods

Design

Data were from a prospective, longitudinal, practice-based, nonrandomized comparative study of self-referring patients with chronic and acute LBP treated by doctors of chiropractic (DCs) and primary-care medical doctors (MDs). This comparative study design is considered appropriate for cost-effectiveness analysis, although it does not yield the level of evidence of a randomized trial. The study enrolled 2872 patients over a 2-year period (1994-1996) from the practices of 60 DCs and 111 MDs in 51 DC and 14 general practice community clinics. Except for one medical clinic located in Vancouver, Washington, all medical and chiropractic clinics were located in Oregon. Patient data were obtained through self-administered questionnaires at the initial visit and mailed follow-up questionnaires. Practitioners were not asked to alter their usual management of LBP for the study.

Participants

Patients with the primary complaint of acute or chronic LBP were eligible to participate if they were at least 18 years old, ambulatory, and English literate. Pain had to be of mechanical origin (i.e., not due to tumors, inflammatory disease, or organic referred pain). Patients were excluded if they had received care from a provider of the same type as the enrolling clinician within the previous 6 weeks, were pregnant, or had contraindications to spinal manipulation. All participants signed a consent form that explained the study and the participant’s rights. The study was approved for protection of human subjects by the Western States Chiropractic College Institutional Review Board.

Treatment

The study clinicians provided a variety of health services. The salient features of chiropractic care were spinal manipulation, physical modalities, exercise plan, and self-care education. Medical patients received prescription drugs, exercise plan, and self-care advice; approximately 25% were referred for physical therapy.
Outcome and Baseline Measures

Information collected at the baseline included history of LBP before the baseline episode, duration and severity of current episode, as well as comorbidities (arthritis, respiratory conditions, gastrointestinal problems, gynecologic problems, hypertension, and other chronic conditions), physical and mental health status, demographics, insurance characteristics, confidence in successful treatment outcome, and a depression screen. Severity of pain and disability were measured 7 times after the baseline visit, only two of which are included in this report. Physical/mental health and patient satisfaction were measured at 12 months. Clinical and satisfaction outcomes were evaluated on 100-point scales. Pain severity, a primary clinical outcome, was measured on a 100-mm Visual Analogue Scale (VAS): “no pain” (0) to “excruciating pain” (100). The VAS is a commonly used, validated pain measure. Functional disability, the other primary clinical outcome, was measured with the Revised Oswestry Disability Questionnaire, a 10-item, 100-point scale assessing pain and daily activities. A higher score on this valid and responsive instrument indicates greater disability. Physical and mental health were evaluated with subscales of the Short Form (SF)-12 questionnaire, a validated short version of the Medical Outcomes Study SF-36. A 3-item depression questionnaire appended to the SF-12 was used to screen for major depression/dysthymia. Two questions measured trust of the provider types, and one question evaluated confidence in treatment success. These 3 were measured on 6-point Likert scales dichotomized for the analysis. Chronic LBP was defined as an episode of at least 7 weeks duration at enrollment. Patient data were obtained using self-administered questionnaires.

Provider practice activities and referrals used in the cost analysis were identified by chart audit for a period of 12 months after baseline. The computation of office-based costs, including x-ray and prescribed medication, have been described elsewhere. Estimates of office costs were based on Medicare/ChiroCode relative value units and Medicare conversion factors. This methodology, increasingly common in economic analyses, provides a standardized measure of costs that does not depend either on the charges, which often do not reflect transaction prices or on the specific amounts collected by the providers in the study.

Estimated total costs for this study included office-based costs plus the estimated costs of advanced imaging, surgical consultation, and referrals to physical therapists. We imputed $600 for advanced imaging costs using data found in Mosely. Our study did not permit us to determine the actual services patients received when referred. We therefore imputed $450 for evaluation by a surgeon to any patient with one or more surgical referrals. This was based on charges data per claimant found in Mushinski, adjusted for the proportion of provider charges that are actually reimbursed. We also imputed $220 to any patient with one or more referrals to a physical therapist, based on Cherkin et al. All costs are in constant 1995 US dollars.

Statistical Analysis

The analysis consisted of forced-entry, multiple regression models conducted separately for each cost and clinical outcome at 3 and 12 months after the initial study visit. Acute and chronic LBP were analyzed separately because of the long recognized distinction between these conditions; 2780 patients who could be identified as acute or chronic were included in the analysis. We examined the impact of provider type on total costs (primary cost analysis) and office costs defined above. The effects of provider type on the primary clinical outcomes, pain and disability, have been reported for all follow-up. Summary scores for patient satisfaction and improvement in physical and mental health at 12 months were secondary outcomes not analyzed previously.

The effects of provider type were adjusted for all independent variables in the models. The variables entered in the models were selected a priori based on general interest in research studies (eg, age and sex) or because they have been previously reported to affect low back outcomes. An additional variable was added to help control for desirability of physician type. This consisted of the difference in trust in chiropractors and MDs, measured on 6-point Likert scales, that we found to be predictive of choice of type of doctor. For clinical outcomes, independent variables consisted of baseline severity, LBP history, referred pain above knee, referred pain below knee, depression, comorbidity, sex, age, smoking, a measure of relative desirability of care type, and interaction effects. Independent variables for cost analysis additionally included variables that were not found previously to be predictors of clinical outcomes: health insurance, marital status, and income. The incremental cost of additional clinical improvements associated with treatment by chiropractors rather than MDs was then computed.

As a secondary analysis, a natural log transformation was applied to total and office cost variables used in the regressions to take into account skewness of these variables. Incremental log costs and associated cost ratios were computed.

Statistical significance was set as \( P < .01 \), and a clinical important difference between groups for the primary outcomes was set at 10 points a priori. Analyses were performed using SAS Version 8.2 (SAS Institute Inc, Cary, NC).

Results

Response rates for the clinical outcomes questionnaires were 66.0% at 3 months and 62.6% at 12 months; these were uniform across groups. Sensitivity analyses revealed no effect of missing data on adjusted group differences. There were very small differences in primary outcomes.
between the results from the subsample of patients with complete data over 4 years and the entire sample. In addition, predictive models showed no effect of missing data on the primary outcomes at 12 months. Complete data for all variables included in cost analyses were available for 38% of chronic and 50% of acute patients. Most data were available for almost all patients, so we were able to accurately profile costs incurred by those excluded because of missing data. The costs incurred by such patients differed little from costs of patients with complete data.

**Patient Characteristics**

The demographic, payment, complaint, general health, and psychosocial characteristics for the 4 cohorts are presented in Table 1. Most differences between MD and DC cohorts were statistically significant. However, only a few of these differences were clinically important and emerged as predictors of clinical or cost outcomes. For chronic patients, MD patients had greater disability, poorer physical health, and greater prevalence of pain radiating below the knee. For the acute cohorts, less than 10% of MD patients and more than 40% of DC patients paid for care out of pocket.

**Cost Outcomes**

Table 2 summarizes unadjusted costs. The impact of the inclusion of costs incurred outside clinicians’ offices on the costliness of MD and DC treatment is notable. Patients
treated by DCs were referred to outside providers infrequently. As a result, mean total costs for DC patients were a little higher than office costs ($6-$10 at 3 months and $10-$14 at 12 months). On the other hand, for patients treated by MDs, referral and advanced imaging accounted for a large fraction of mean total costs (acute, 24%-36% or $51-$105; chronic, 48%-50% or $63-$135). Office costs for DC care were 78% to 82% higher than MD care for acute patients and 52% to 60% higher for chronic patients. In contrast, total costs of DC care were only 22% greater than MD care for acute patients and 16% less than MD care for chronic patients.

Table 3 reports adjusted differences in costs and outcomes. Office costs for chiropractic treatment had higher

<table>
<thead>
<tr>
<th>Table 2. Cost and clinical outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>-----------------------------------</td>
</tr>
<tr>
<td><strong>Chronic DC</strong></td>
</tr>
<tr>
<td>3 months</td>
</tr>
<tr>
<td>Office costs $174 $195</td>
</tr>
<tr>
<td>Median $104 $103</td>
</tr>
<tr>
<td>Total costs $180 $209</td>
</tr>
<tr>
<td>Median $108 $103</td>
</tr>
<tr>
<td>Pain 22.8 25.4</td>
</tr>
<tr>
<td>Disability 15.3 16.1</td>
</tr>
<tr>
<td>12 mo</td>
</tr>
<tr>
<td>Office costs $222 $288</td>
</tr>
<tr>
<td>Median $116 $103</td>
</tr>
<tr>
<td>Total costs $232 $311</td>
</tr>
<tr>
<td>Median $123 $135</td>
</tr>
<tr>
<td>Pain 23.9 27.0</td>
</tr>
<tr>
<td>Disability 16.1 17.1</td>
</tr>
<tr>
<td>Physical health 14.7 18.3</td>
</tr>
<tr>
<td>Mental health 4.9 20.5</td>
</tr>
<tr>
<td>Satisfaction 86.4 19.9</td>
</tr>
<tr>
<td><strong>Chronic MD</strong></td>
</tr>
<tr>
<td>3 months</td>
</tr>
<tr>
<td>Office costs $107 $75</td>
</tr>
<tr>
<td>Median $84 $84</td>
</tr>
<tr>
<td>Total costs $212 $253</td>
</tr>
<tr>
<td>Median $103 $103</td>
</tr>
<tr>
<td>Pain 16.7 29.9</td>
</tr>
<tr>
<td>Disability 28.3 20.4</td>
</tr>
<tr>
<td>12 mo</td>
</tr>
<tr>
<td>Office costs $146 $153</td>
</tr>
<tr>
<td>Median $103 $103</td>
</tr>
<tr>
<td>Total costs $281 $355</td>
</tr>
<tr>
<td>Median $135 $135</td>
</tr>
<tr>
<td>Pain 18.9 31.8</td>
</tr>
<tr>
<td>Disability 29.4 20.6</td>
</tr>
<tr>
<td>Physical health 15.8 20.8</td>
</tr>
<tr>
<td>Mental health 4.9 19.5</td>
</tr>
<tr>
<td>Satisfaction 71.3 22.7</td>
</tr>
<tr>
<td><strong>Acute DC</strong></td>
</tr>
<tr>
<td>3 months</td>
</tr>
<tr>
<td>Office costs $161 $183</td>
</tr>
<tr>
<td>Median $101 $102</td>
</tr>
<tr>
<td>Total costs $171 $202</td>
</tr>
<tr>
<td>Median $102 $70</td>
</tr>
<tr>
<td>Pain 39.9 27.3</td>
</tr>
<tr>
<td>Disability 41.8 30.3</td>
</tr>
<tr>
<td>12 mo</td>
</tr>
<tr>
<td>Office costs $206 $284</td>
</tr>
<tr>
<td>Median $121 $82</td>
</tr>
<tr>
<td>Total costs $218 $305</td>
</tr>
<tr>
<td>Median $124 $89</td>
</tr>
<tr>
<td>Pain 41.9 28.5</td>
</tr>
<tr>
<td>Disability 41.9 28.5</td>
</tr>
<tr>
<td>Physical health 20.5 19.9</td>
</tr>
<tr>
<td>Mental health 4.9 18.9</td>
</tr>
<tr>
<td>Satisfaction 90.2 16.4</td>
</tr>
<tr>
<td><strong>Acute MD</strong></td>
</tr>
<tr>
<td>3 months</td>
</tr>
<tr>
<td>Office costs $90 $66</td>
</tr>
<tr>
<td>Median $69 $69</td>
</tr>
<tr>
<td>Total costs $141 $183</td>
</tr>
<tr>
<td>Median $70 $70</td>
</tr>
<tr>
<td>Pain 41.8 30.3</td>
</tr>
<tr>
<td>Disability 41.8 30.3</td>
</tr>
<tr>
<td>12 mo</td>
</tr>
<tr>
<td>Office costs $113 $117</td>
</tr>
<tr>
<td>Median $82 $82</td>
</tr>
<tr>
<td>Total costs $176 $245</td>
</tr>
<tr>
<td>Median $89 $89</td>
</tr>
<tr>
<td>Pain 41.9 28.5</td>
</tr>
<tr>
<td>Disability 41.9 28.5</td>
</tr>
<tr>
<td>Physical health 20.5 19.9</td>
</tr>
<tr>
<td>Mental health 4.9 18.9</td>
</tr>
<tr>
<td>Satisfaction 76.0 22.6</td>
</tr>
</tbody>
</table>

All clinical outcomes were normalized to a 100-point scale. A higher value denotes greater satisfaction or greater improvement in pain, disability, physical health, and mental health. All improvement scores were statistically significant ($P < .01$).

<table>
<thead>
<tr>
<th>Table 3. Adjusted mean differences (DC-MD) in costs and outcomes improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>---------------------------------</td>
</tr>
<tr>
<td><strong>Chronic patients</strong></td>
</tr>
<tr>
<td>3 months</td>
</tr>
<tr>
<td>Office costs $142 $37 $0.000</td>
</tr>
<tr>
<td>Office costs (log) 0.69 0.22 0.002</td>
</tr>
<tr>
<td>Total costs $5 $52 $0.931</td>
</tr>
<tr>
<td>Total costs (log) 0.22 0.25 $0.39</td>
</tr>
<tr>
<td>Pain 10.5 2.0 $0.000 $13.5 $0.4</td>
</tr>
<tr>
<td>Disability 8.8 1.6 $0.000 $16.1 $0.5</td>
</tr>
<tr>
<td>12 months</td>
</tr>
<tr>
<td>Office costs $158 $60 0.009</td>
</tr>
<tr>
<td>Office costs (log) 0.58 0.23 $0.14</td>
</tr>
<tr>
<td>Total costs $1 $80 $0.993</td>
</tr>
<tr>
<td>Total costs (log) 0.10 0.26 $0.715</td>
</tr>
<tr>
<td>Pain 7.3 2.1 $0.000 $21.6 $0.1</td>
</tr>
<tr>
<td>Disability 5.4 1.7 $0.000 $29.2 $0.1</td>
</tr>
<tr>
<td>Physical health 3.0 3.6 $0.396 $52.2 $0.2</td>
</tr>
<tr>
<td>Mental health 1.2 3.7 $0.757 $136.4 $0.7</td>
</tr>
<tr>
<td>Satisfaction 18.1 4.9 $0.000 $8.7 $0.0</td>
</tr>
<tr>
<td><strong>Acute patients</strong></td>
</tr>
<tr>
<td>3 months</td>
</tr>
<tr>
<td>Office costs $93 $25 $0.000</td>
</tr>
<tr>
<td>Office costs (log) 0.48 0.15 $0.002</td>
</tr>
<tr>
<td>Total costs $42 $35 $0.224</td>
</tr>
<tr>
<td>Total costs (log) 0.18 0.17 $0.288</td>
</tr>
<tr>
<td>Pain 3.6 1.3 $0.005 $25.7 $11.7</td>
</tr>
<tr>
<td>Disability 3.9 1.1 $0.000 $23.8 $10.8</td>
</tr>
<tr>
<td>12 months</td>
</tr>
<tr>
<td>Office costs $112 $38 $0.003</td>
</tr>
<tr>
<td>Office costs (log) 0.39 0.16 $0.017</td>
</tr>
<tr>
<td>Total costs $43 $47 $0.352</td>
</tr>
<tr>
<td>Total costs (log) 0.13 0.18 $0.453</td>
</tr>
<tr>
<td>Pain 9.2 2.5 $0.000 $12.2 $4.7</td>
</tr>
<tr>
<td>Disability 2.7 1.1 $0.012 $41.7 $16.1</td>
</tr>
<tr>
<td>Physical health 5.4 2.5 $0.032 $20.8 $8.0</td>
</tr>
<tr>
<td>Mental health 14.0 3.1 $0.000 $8.0 $3.1</td>
</tr>
</tbody>
</table>

Adjusted mean differences between DC and MD are the predicted mean differences from the regression models. Positive values indicate greater cost or greater improvement in outcomes for DC patients. CER indicates the incremental cost-effectiveness ratios: adjusted mean difference in cost divided by adjusted mean difference in outcomes. Office costs were used in the numerator of CER1, and total costs were used in the numerator of CER2.
costs for both chronic and acute patients at the 3- and 12-month intervals ($93-$158, \( P < .01 \)). However, when costs of advanced imaging and referrals were included (primary analysis), costs of DC treatment were not significantly different from those of medical treatment at either the 3-month or the 12-month interval. Adjusted differences were $5 and $1 at the two intervals for chronic patients (\( P > .90 \)) and $42 and $43 for acute patients (\( P > .20 \)). The impact of chiropractic treatment on costs remained unchanged when a log transform of costs was used in the analysis. Adjusted DC office costs were 1.5 to 2.0 times greater (\( P < .01 \)), whereas DC total costs were only 1.1 to 1.2 times greater and not statistically significant (\( P > .25 \)).

The regression models not only adjusted outcomes for group differences in the independent variables listed under statistical analysis above, but also identified the contribution of predictor variables to the outcomes. The large volume of data necessitates that these results be published elsewhere.

Clinical Outcomes

Table 2 shows clinically important and statistically significant, within-group improvement in pain, functional disability, and general health outcomes for all 4 patient cohorts. Patient satisfaction can be considered high for DC patients and somewhat more moderate for MD patients.

Improvement in the pain and disability (primary) outcomes was significantly greater for DC care in both acute and chronic patients. Adjusted mean differences (AMD) in these outcomes were clinically important for chronic patients at 3 months (AMD, 10.5 and 8.8, \( P < .0005 \)). The advantage for DC care in acute patients was small at both 3 and 12 months (AMD < 4, \( P < .01 \)). There was little difference in improvement between DC and MD patients in physical and mental health. One exception was physical health in acute patients (AMD, 9.2; \( P < .0005 \)). Patient satisfaction favored DC care for acute and chronic patients (AMD, 14-18; \( P < .0005 \)).

Adjusted Incremental Cost-Effectiveness Ratios

The additional costs per unit advantage in outcomes for DC care are presented in Table 3. Of note, ratios computed for office costs alone were considerably higher than ratios computed for total costs. For chronic patients, the total cost ratios ranged from approximately $0.1 to $0.5 per point advantage. Specifically, for the primary outcomes at 3 months, there was a $5 additional cost for a 10.5-point advantage in pain and an 8.8-point advantage in improvement. At 12 months, there was only a $1 additional cost but for more modest 7.3- and 5.4-point improvements in these outcomes. For acute patients, the cost ratios were between $24 and $25 per point at 3 months and $8 to $42 per point at 12 months. The cost ratios reflect greater cost and smaller advantage in primary outcomes than for chronic patients.

Also notable are the small ratios for large differential satisfaction in both acute and chronic patients.

Discussion

Back pain is experienced by 80% of adults during their lives\(^a\) and accounts for 2.5% of US health care expenditures.\(^b\) Arguably, the relative cost-effectiveness of medical and chiropractic care is an urgent economic and health policy issue, one for which evidence is especially limited. Much of the recent work on cost-effectiveness has been conducted abroad.\(^{25-27}\) With cost structures in the United States that are very different from other countries,\(^{48,49}\) our work fills important information gaps that can help with policy and health plan decisions. We include a broad set of outcomes indicators as well as comprehensive cost data for large samples of patients. Furthermore, we have been able to adjust both costs and outcomes for a variety of confounding factors to provide clear relative cost indicators.

Our study had several important findings. First, office costs alone are not appropriate outcomes for a comparison of medical and chiropractic care. Medical office costs do not include physical therapy, whereas physical modalities are usually performed in chiropractic offices.\(^{29}\) These and other referral costs (advanced imaging and other provider care) appear to be the great equalizers for medical and chiropractic care. The appropriateness of advanced imaging and referral were not investigated in this study. Clearly, over- and underuse could have a dramatic effect on relative cost-effectiveness.

Chiropractic appears relatively cost-effective compared with medical care for the treatment of chronic LBP in pain and functional disability improvement. This was evidenced by a relative clinical benefit, particularly in the short term, concomitant with no difference in total costs. The picture for acute patients is somewhat less clear. There was only a small advantage for chiropractic care in outcomes with additional but statistically insignificant costs.

Two recent randomized trials addressed cost-effectiveness of manipulation/chiropractic care. Using a formal analysis, a trial in the United Kingdom found that manipulation is cost-effective for back pain.\(^{25}\) Kominski et al\(^{50}\) found, at an 18-month follow-up, that chiropractic care was more expensive than medical care, but chiropractic care with physical modalities was less expensive than medical care with physical therapy. Outcomes were comparable across the 4 groups. This study supports our contention that ancillary care such as physical modalities need to be considered in cost-effectiveness studies. The absence of group differences in outcomes at 18 months is consistent with our study findings reported previously; chiropractic and medical care differences vanished between 12 and 24 months.\(^{15}\)

Although most cost comparisons have been favorable to chiropractic, several studies for the United States have
reported that chiropractic care costs more than treatment provided by primary care physicians.\textsuperscript{19,20} For example, general practitioners had the lowest charges over episodes of care, with DCs and orthopedists the highest, in a study using 1974 to 1982 data from the RAND Health Insurance Experiment.\textsuperscript{19}

In particular, our findings were in contrast to the seminal, nonrandomized comparative study by Carey et al.\textsuperscript{20} who found equivalent outcomes but the highest costs for urban DCs and orthopedists and the lowest for primary care and health maintenance organizations. However, their cost data reflected charges rather than payments, which are often much lower than charges. Their costs were also evaluated for a single episode, rather than a fixed period. Many investigators believe that the episode is the appropriate unit of analysis.\textsuperscript{51} However, costs over a fixed period capture recurrences and, thus, may be the more practical approach from the perspective of payers and policy makers.

Our results were consistent with Carey et al\textsuperscript{20} and a trial by Cherkin et al\textsuperscript{21} in finding greater satisfaction with chiropractic care than with other interventions. We do not know how to value satisfaction against costs at this time but feel that satisfaction is an outcome that merits consideration in cost-effectiveness studies.

The RAND\textsuperscript{19} study provides an example of cost-minimization analysis, a method that is,\textsuperscript{31} “appropriate if the alternatives have identical consequences” including “side effects and adverse events.” Despite these caveats, cost minimization has been the dominant methodology used in US cost analyses. In a subsequent example, patients with back and neck pain treated by chiropractors in one health maintenance organization had lower costs than those treated by other providers.\textsuperscript{16} The authors recognized that they did not control for differences in comorbidities, chronic illnesses, or severity but only inferred from other data that there were no substantial differences in underlying illnesses.

A more widely cited study applied an incremental spending methodology to a large database of fee-for-service patients with LBP.\textsuperscript{17,18} Chiropractic users had far lower outpatient and total costs for their episodes of care than nonusers. Although the analysis included controls for differences in patients’ insurance and sociodemographic characteristics, controls for the severity of the condition and health status of the patient were limited. The study also did not include any patient outcomes measures. In the large managed care network study in California, where members with chiropractic coverage showed lower annual health care expenditures and lower use rates per episode of back pain than those without chiropractic coverage,\textsuperscript{24} there were no patient outcomes measures that could lead to stronger evidence of chiropractic’s relative cost-effectiveness. Our contribution examined both costs and outcomes to report results through easily understood incremental cost-effectiveness ratios.

Nevertheless, several limitations may have affected the study outcomes and generalization of findings. It is well-known that observational studies are more susceptible to bias than randomized controlled trials from unknown factors associated with patients and providers. Control for relevant confounding variables would have the greatest validity in inferring that the costs and outcomes are not attributable to other extraneous factors in observational studies.\textsuperscript{32,33} Our study statistically controls for a broad set of potentially confounding variables to evaluate cost and effectiveness in actual practice when patients can select the providers of their choice. A well-designed observational study can thus overcome a major weakness of randomized trials, their artificial design and limited generalization to clinical practice.\textsuperscript{52} Only large, pragmatic, randomized trials that do not control patient management can yield more accurate estimates of adjusted cost and outcomes differences between medical and chiropractic care.

Hospitalization/surgical costs were not available for our analysis. Because there was a greater referral rate for surgical evaluation from MDs and the hospitalization rate is known to be higher for medical patients,\textsuperscript{19} it is likely that inclusion of hospitalization/surgery would have increased medical costs disproportionately.

Over-the-counter (OTC) drug costs were also excluded from the analysis. We found OTC drug costs difficult to estimate, because the data collected did not account for the large variation in drug type and pill dosage. Drug costs appeared to be relatively small compared to provider costs, so bias was probably small. It is unknown whether there was differential consumption of OTCs between chiropractic and medical patients.

Caution must be taken in generalizing study findings from a regional study to national practice. Chiropractic scope of practice varies from state to state,\textsuperscript{53} permitting different modalities for the treatment of LBP. For example, Oregon’s scope of practice included physical modalities, whereas neighboring Washington’s did not. Caution must also be used in light of the continual evolution in health care financing and reimbursement mechanisms. The study controlled for some differences in patients’ insurance characteristics, and these results will be reported elsewhere. However, the study design, conceived in the early 1990s, did not anticipate the extent of the shift toward managed care or of other developments such as consumer-driven health plans.

**Conclusions**

This study supports the generalizability of systematic reviews of the efficacy of spinal manipulation for pain and functional disability to the effectiveness of chiropractic care in clinical practice. Our findings are consistent with the review findings that spinal manipulation–centered therapy is
as least as good as, and in some cases, better than other treatments of LBP.8-10 Although randomized trials found an advantage for chiropractic care in costs, our study leaned toward comparability.

Chiropractic patients with chronic LBP showed an advantage over medical patients in pain, disability, and satisfaction outcomes without additional costs. Chronic pain and disability outcomes were clinically important in the short term and of lesser magnitude in the long term. Satisfaction with chiropractic care was considerably greater for both acute and chronic patients at both time points. Although the advantages in pain and disability were small for acute patients with LBP, it is important to consider that these gains can be obtained with, at most, small increased costs. With their mission to increase value and respond to patient preferences, health care organizations and policy makers need to reevaluate the appropriateness of chiropractic as a treatment option for LBP.

**Practical Application**

- Chiropractic care is relatively cost-effective compared with primary medical care for the treatment of chronic LBP, particularly in the short term.
- Chiropractic and medical care are comparable in cost and effectiveness for acute LBP.
- Healthcare organizations and policy makers should consider the appropriateness of chiropractic as a treatment option for LBP.

**References**

25. UK BEAM Trial Team. United Kingdom back pain exercise and manipulation (UK BEAM) randomised trial: cost


EFFECTS OF A MANAGED CHIROPRACTIC BENEFIT ON THE USE OF SPECIFIC DIAGNOSTIC AND THERAPEUTIC PROCEDURES IN THE TREATMENT OF LOW BACK AND NECK PAIN

Craig F. Nelson, DC, MS, R. Douglas Metz, DC, and Thomas LaBrot, DC

ABSTRACT

Objective: The aim of this study was to measure the effects of a managed chiropractic benefit on the rates of specific diagnostic and therapeutic procedures for the treatment of back pain and neck pain.

Design: This study is a retrospective analysis of claims data from a managed-care health plan over a 4-year period. The use rates of advanced imaging, surgery, inpatient care, and plain-film radiographs were compared between employer groups with and without a chiropractic benefit.

Results: For patients with low back pain, the use rates of all 4 studied procedures were lower in the group with chiropractic coverage. On a per-episode basis, the rates in the group with coverage were reduced by the following: surgery (−32.1%); computed tomography (CT)/magnetic resonance imaging (MRI) (−37.2%); plain-film radiography (−23.1%); and inpatient care (−40.1%). On a per-patient basis, the rates were reduced by the following: surgery (−13.7%); CT/MRI (−20.3%); plain-film radiography (−2.2%); and inpatient care (−24.8%). For patients with neck pain, the use rates were reduced per episode in the group with chiropractic coverage as follows: surgery (−49.4%); CT/MRI (−45.6%); plain-film radiography (−36.0%); and inpatient care (−49.5%). Per patient, the rates were surgery (−31.1%); CT/MRI (−25.7%); plain-film radiography (−12.5%); and inpatient care (31.1%). All group differences were statistically significant.

Conclusion: For the treatment of low back and neck pain, the inclusion of a chiropractic benefit resulted in a reduction in the rates of surgery, advanced imaging, inpatient care, and plain-film radiographs. This effect was greater on a per-episode basis than on a per-patient basis. (J Manipulative Physiol Ther 2005;28:564-569)

Key Indexing Terms: Chiropractic Benefit; Low Back Pain; Neck Pain

A growing body of clinical evidence and expert opinion indicates that a more conservative approach to the treatment of low back pain and cervical spine pain is appropriate.\(^1\)\(^-\)\(^7\) The evidence indicates that procedures and practices such as inpatient care, advanced imaging, surgery, and even plain-film radiographs are only infrequently necessary for the successful treatment of most cases of low back pain and neck pain.\(^8\)\(^-\)\(^17\)

In spite of the evidence, these procedures and practices continue to be used at rates in excess of that which the published literature defines as clinically indicated.\(^18\)\(^-\)\(^22\) With the exception of plain-film radiographs, chiropractors do not directly administer any of these invasive and high-cost procedures, and the rates of those procedures under chiropractic care is essentially zero. However, there are no published data to date that indicate whether the use of chiropractic care affects the overall use rate of these procedures in a given population. This study evaluates the effects of the presence of a chiropractic benefit (and thus, the increased use of chiropractic management) on the rates of these procedures in the treatment of low back pain and neck pain.

METHODS

Using administrative data, this study evaluates the treatment of patients with low back pain and neck pain conditions who are enrolled in a managed-care health plan.
The study period covers April 1, 1997, through March 30, 2001. Individual employers had the option of selecting a health plan medical benefit with or without the addition of a benefit for chiropractic care. Individual health plan members did not make this selection. The study compares those employer groups with and without a chiropractic benefit. The study cohorts are thus:

- **Cohort A** Patients enrolled in medical health plans that also have a chiropractic benefit;
- **Cohort A1** Patients enrolled in medical health plans that also have a chiropractic benefit who received treatment for low back pain or neck pain conditions;
- **Cohort B** Patients enrolled in medical health plans that do not have a chiropractic benefit;
- **Cohort B1** Patients in medical health plans that do not have a chiropractic benefit who received treatment for low back pain or neck pain conditions.

**Identification and Definition of Low Back Pain and Neck Pain Patients and Episodes of Care**

The identification of low back pain and neck pain episodes of care was made by the use of *International Classification of Diseases, Ninth Revision (ICD-9)*, codes that are a part of all administrative data (claims data). A total of 32 ICD-9 codes for neck pain and 41 codes for low back pain were identified as representing these two groups of conditions. An expert panel of chiropractors and medical physicians evaluated this list for appropriateness and completeness.

The aggregation of claims into discrete episodes of care was made on the basis of both a “clean period” of 45 days with no claims as well as the diagnostic category that defines the type of episode. Each episode was initiated by one of the low back pain or neck pain codes in the diagnostic list. All services using a back pain or neck pain code and with a maximum gap of 45 days between claims were aggregated into one episode of care. Thus, a new episode was created if a new diagnostic category was used or encounters were separated by more than 45 days. A claim-free 45-day window was applied to the start and end points of the 4-year study period to identify and include members with nontruncated episodes. For any episode that began during this period but extended beyond March 30, 2001, all services related to that episode (within the 45-day limit) were treated as if they fell within the 4-year period. Similarly, any episode that began within 45 days before April 1, 1997, but extended into the 4-year period was considered to have occurred totally outside the study period and was not used in the analysis. The clean period of 45 days is consistent with previous studies using administrative data.23–26

In addition to evaluating use rates on a per-episode basis, these rates were also measured on a per-patient basis. For each individual who had one or more episodes of low back pain or neck pain, all encounters and services from these episodes were aggregated into one low back pain or neck pain unit of analysis. Thus, an individual might be counted multiple times in the episode analysis but only once in the patient analysis. The per-patient rate represents the overall probability that any individual with a low back or neck pain complaint during the study period would receive the procedures under investigation.

**Data Preparation and Merge**

The data preparation included transfer of all relevant claims data from the two different data sources (see hereinafter), loading of the data onto a common server, and filtering by health member continuous enrollment to ready the data for analysis. Data relevant to patient enrollment (insurance coverage information), health service encounters (ie, dates of service, diagnoses, procedures, costs, and so on), and pharmacy claims were loaded onto the server. SAS version 6.12 (SAS Institute, Cary, NC) was used for all data programming. Before the analysis, the data were validated as follows: verify receipt of correct data (names, number of files, number of records contained in each file with each respective data source); validate the format of the data (character, numeric, and length); identify key variables in the data sets (age, diagnosis codes, and others); produce frequency reports of the data and validate the variable’s contents, again working with each respective data source; run algorithms (computer programs designed to detect implausible data) to ensure the integrity of key variables (eg, *International Classification of Diseases [ICD-9]* and *Current Procedural Terminology [CPT-4]* codes); and send reports produced from previous to each respective company for verification.

For patients with chiropractic coverage, there is an entirely separate and distinct treatment and storage of claims data for their chiropractic care than for their medical care. For this study, a patient’s chiropractic claims must be merged with their medical claims producing a single claims file for each covered patient. Merging of the data sets was accomplished using one of the following methods. Each health plan member was assigned a unique identification number that is used for both the medical and chiropractic claims. This number was used to link a patient’s chiropractic claims with their medical claims. Should there be no common unique member identifier (because of data entry errors), the data were linked using both member social security number and date of birth. Once the data were linked, a unique identifier was created, and name, address, and social security number was purged from the data set to assure patient confidentiality. Any data not linked by these two methods were eliminated from the study.

**Data Analysis**

Use rates for the following procedures were measured in the two cohorts: inpatient care for back pain and neck pain;
computed tomography/magnetic resonance imaging use related to back pain and neck pain; surgeries for back pain and neck pain; and plain-film radiographs for back pain and neck pain.

Rates for the procedures were calculated both on a per-episode basis and a per-patient basis. Tests for group differences were done using Wilcoxon test.

**RESULTS**

**Data Merging**

Of the chiropractic claims, data files from April 1, 1997, through March 30, 2001, 98.3% were successfully merged with the managed-care organization’s claims files. The balance (1.7% of claims) was eliminated from the analysis.

**Study Population Characteristics**

An analysis was conducted on those patients who had changed their chiropractic coverage status during calendar year 2000. (The 4-year data contain a slightly greater number of total patients because it also includes those who did change their chiropractic coverage status at some point during the study.) There were small differences in the rates of comorbid conditions and demographic characteristics, and in the study populations. The group with coverage had fewer comorbid conditions in most of the categories studied and was slightly younger. A summary of these findings is shown in Tables 1 and 2.

**Procedure Rates Among Low Back Pain Patients**

Comparing the cohort with chiropractic coverage to cohort without coverage, the rates of surgery, advanced imaging, inpatient care, and plain-film radiographs were all lower in the cohort with chiropractic coverage. On a per-patient basis, the rates were lower in the chiropractic coverage cohort by the following amounts: surgery, 13.7%; advanced imaging, 20.3%; inpatient visits, 24.8%;
and plain-film radiographs, 2.2%. On a per-episode basis, the rates were reduced by the following amounts: surgery, 32.1%; advanced imaging, 37.2%; inpatient visits, 40.1%; and plain-film radiographs, 23.1%. Table 3 shows the actual rates of these procedures in the two cohorts. All differences were statistically significant (P < .01).

**Procedure Rates Among Neck Pain Patients**

Similarly, for neck pain, comparing the cohort with chiropractic coverage to cohort without chiropractic coverage, the rates of surgery, advanced imaging, inpatient care, and plain-film radiographs were all reduced in the cohort with chiropractic coverage. On a per-patient basis, the rates were reduced by the following amounts: surgery, 31.1%; advanced imaging, 25.7%; inpatient visits, 31.1%; and plain-film radiographs, 12.5%. On a per-episode basis, the rates were reduced by the following amounts: surgery, 49.4%; advanced imaging, 45.6%; inpatient visits, 49.5%; and plain-film radiographs, 36.0%. Table 4 shows the rates of these procedures in the two cohorts. All differences were statistically significant (P < .01).

**Discussion**

In all study categories, there were statistically significant reductions in the rates of surgery, advanced imaging, inpatient care, and plain-film radiography. Previously published data from this study showed that chiropractic care was used almost entirely as a substitution for medical care for back pain and neck pain complaints. The clinical treatment of back and neck pain complaints tends to be highly variable. Specifically, the use rates of the procedures here under investigation tend to vary significantly in different care delivery systems, and these rates are not driven solely, or even principally, by clinical variables. Given the amount of discretion that is exercised in the use of these procedures, it is not surprising to find that when the option of accessing chiropractic care is provided, it would produce the reductions seen in this study.

The difference in rates is greater as measured on a per-episode basis than on a per-patient basis. Also, as demonstrated in a previous publication, chiropractic care (which is only present in cohort A1) tends to generate more episodes of care per patient than medical care and thus artificially further reduces the rate per episode. The per-patient comparison is the more valid and meaningful because it captures the actual probability of an individual patient’s likelihood of receiving any of the specified interventions.

It is also notable that the rates of plain-film radiographs are lower in the group with chiropractic coverage. Unlike the other 3 procedure categories, chiropractors themselves provide plain-film radiographs. Previous studies have shown that chiropractors may use plain-film radiographs at a higher rate than medical physicians, and thus, the substitution phenomenon might produce an increase rather than a decrease in these rates. The reductions seen here are attributable to specific treatment policies designed to reduce the rates of use of radiographs in chiropractic care and may not reflect the rate of use in an unmanaged system.

Given the significant differences in the rates of procedures found in this study, it is important to consider which rates represent a more optimal level of care for back pain and neck pain. These differences exist within the same geographic area and between reasonably similar patient groups. It is a plausible assumption that if the same standard of care were to be applied to these two groups, very similar use rates would result. However, it is not possible to answer this question directly because there are simply no clinical outcomes measured in this study.

There are also no standard benchmarks for use rates against which these rates can be compared. However, there is a large body of scientific literature and guidelines on the treatment of back pain that provides some insight into this question. This body of evidence coalesces around 3 recurrent themes: (1) most back pain episodes are categorized as “mechanical” or “uncomplicated” and do not require aggressive interventions, (2) existing guidelines specify that less aggressive and invasive diagnostic and therapeutic treatment of back pain will lead to better clinical outcomes, and (3) there is poor compliance with these guidelines and departures from the guidelines are most likely in the direction of overuse of invasive procedures. Given these observations, it is probable that the rates of use of these procedures are greater than optimal in both groups and any change in the direction of decreased use may result in positive effects on health-care costs, outcomes, and patient safety. The literature on neck pain is less comprehensive than that on low back pain, but it also points in the same direction. The conclusion that the presence of a chiropractic benefit results in more appropriate use diagnostic and therapeutic procedures is suggested but not proven by this study.

The cohort with chiropractic coverage was slightly younger with slightly fewer comorbidities than the cohort without chiropractic coverage. These differences may have contributed to the reductions in procedure rates. However, the magnitudes of these demographic and clinical differences are quite small compared with the magnitudes of the reductions, and it is unlikely that they account for more than a small percentage of the changes.

The results of this study cannot be generalized beyond the specific health care systems involved in the study. It is well understood that the use rates of specific diagnostic and therapeutic procedures are highly dependent upon variables such as local practice habits, economic incentives, and other health plan characteristics. A different set of these variables would undoubtedly produce a different result. However, the direction of these results (a reduction in rates among those with chiropractic coverage) might be more robust relative to these variables. As long as
chiropractic care is being used as a substitute for medical care and as long as chiropractors do not directly provide these procedures (with the exception of plain-film radiographs), it is likely that a reduction in advanced imaging, surgery, and inpatient care would be seen. Finally, there are no clinical outcomes measured in this investigation. No inferences can be made on the relative clinical benefits of the different procedure rates except as suggested by the existing literature on this subject.

CONCLUSION

Among employer groups with chiropractic coverage compared with those without such coverage, there is a significant reduction in the use of high-cost and invasive procedures for the treatment of low back pain and neck pain. The presumed mechanism of this effect is the substitution of chiropractic care for medical care for the treatment of back and neck pain. The resultant chiropractic care is far less likely to lead to the use of these invasive procedures. This reduction is more pronounced when measured on a per-episode basis than on a per-patient basis.

REFERENCES

31. Freeborn D, Shye D, Mullooly JP, Eraker S, Romeo J. Primary care physicians’ use of lumbar spine imaging tests: effects of


A CROSS-SECTIONAL STUDY COMPARING PAIN AND DISABILITY LEVELS IN PATIENTS WITH LOW BACK PAIN WITH AND WITHOUT TRANSITIONAL LUMBOSACRAL VERTEBRAE

Cynthia K. Peterson, DC, MMedEd, Jennifer Bolton, PhD, MSc, William Hsu, DC, and Angela Wood, DCR(R)

ABSTRACT

Objective: To determine whether patients with transitional lumbosacral vertebrae report more pain and disability compared with patients with normal lumbar vertebrae.

Methods: Radiographic and questionnaire data were collected from 353 patients with low back pain. Back pain severity was measured using 2 scales: one for pain over the entire episode and the other for pain during the previous week. All patients completed the Revised Oswestry Disability Questionnaire before radiography was performed. Patients were divided into 2 groups: those with and those without a transitional lumbosacral vertebra. Differences between patient groups were investigated using the unpaired t test. Multiple linear regression analysis was applied to investigate the effect of the transitional lumbosacral vertebrae on pain and disability controlling for the effects of age and sex.

Results: Forty-three patients (12.2%) had a transitional lumbosacral vertebra. There were no differences in pain or disability levels between the 2 groups on any of the pain scales or Revised Oswestry subscales. Older patients reported significantly more pain ($P = .039$) and disability ($P = .002$) than younger patients.

Conclusions: The presence of a transitional lumbosacral vertebra in this group of patients was not related to an increased level of reported low back pain or disability. (J Manipulative Physiol Ther 2005;28:570-574)

Key Indexing Terms: Lumbar Vertebrae; Sacrum; Musculoskeletal Abnormalities; Pain; Transitional Lumbosacral Vertebrae

Approximately 4% to 8% of the population have transitional vertebrae at the lumbosacral junction,1 making them commonly seen radiographic anomalies. These transitional lumbosacral segments were formerly known as either “lumbarization” or “sacralization,” reflecting the assumption that either the first sacral segment showed morphologic features of a lumbar vertebra or the last lumbar vertebra became partially or completely incorporated into the sacrum. Various categories of these anomalies were developed, depending on the presence and location of accessory joints or fusions,2-4 with type II of Castellvi being most common.5 Recent literature has critiqued the use of these older terms sacralization/lumbarization due to the very poor inter- and intraexaminer reliability of the use of these labels,6 as well as the fact that it is considered within normal limits for patients to have one extra or one less lumbar vertebra. In addition, anomalies in one transitional region of the spine are often associated with anomalies at other transitional junctions.2

Most transitional lumbosacral vertebrae (TLSV) are asymmetric in morphology, often demonstrating unilateral accessory joints or fusion.7,7 Because of this asymmetry in shape, clinicians have inferred asymmetry in motion segment function, leading to alteration of stresses, particularly at the motion segment immediately above the transitional anomaly.8 Indeed, several studies have confirmed that patients with TLSV have earlier and more severe degeneration at the motion segment immediately above this anomaly,7,9-11 irrespective of the type (category) of lumbosacral segment visualized. It is hypothesized that this degeneration...
is caused by the relative lack of motion between the transitional segment and the segment below, resulting in hypermobility at the motion segment immediately above the anomaly.

Several authors have suggested that patients with this congenital anomaly have more low back pain (LBP) compared with patients without transitional segments. Although studies have shown that patients with TLSV do not have a higher incidence of disk herniation, overall, the location of an herniation, when it does occur, is much more commonly found at the motion segment immediately above the anomaly. This again is most likely due to the increased motion at this level and resultant stresses on the disk.

When abnormal radiographic findings are detected, it is tempting to attribute the patients’ symptoms to the x-ray findings, whether there is research evidence to support these conclusions. Studies investigating the link between degenerative changes in the lumbar spine and pain have found conflicting results with only one study exploring the relationship between degeneration and functional disability, as well as including the facet articulations and the intervertebral disks in the data collection. The weight of the evidence indicates that patients with LBP with lumbar spine degeneration do not have more disability than patients without degeneration, other than a slight increase in pain while in the standing position.

Although people with TLSV tend to have more severe and earlier degeneration at the motion segment above the anomaly, it does not necessarily follow that they should have more pain or disability simply because of this degeneration. If indeed there is more pain associated with transitional anomalies, it may be due to factors other than degeneration. A recent study from Finland compared only men with and without transitional lumbosacral vertebrae for differences in pain levels and degeneration using magnetic resonance imaging. Although they again found an increased risk of degeneration at the disk above the anomalous level, they found no association between these anomalies and LBP. In addition, the link between TLSV and disability has never been investigated. Therefore, this cross-sectional study was undertaken to determine if both male and female patients with these anomalies present with differences in their pain and disability levels, compared with patients with “normal” lumbosacral morphology.

**Methods**

Data were collected from consecutive patients with LBP, referred to the Anglo-European College of Chiropractic and the Canadian Memorial Chiropractic Clinic for radiographic examination of the lumbar spine as part of their clinical examination. The study was approved by the institutional review boards at both institutions, and all subjects provided informed consent. Exclusion criteria included patients younger than 17 years and those with radiographic evidence of spondylolytic spondylolisthesis, congenital blocked vertebrae, sagittal cleft vertebrae, malignant bone tumors, pain-producing benign bone tumors, acute fractures, spinal infection, or avascular necrosis of the hip. The exclusion criteria were applied upon radiographic interpretation. No comparison group of patients without LBP could be included because of the ethics involved in exposing asymptomatic individuals to ionizing radiation.

**Questionnaires**

While waiting for their radiographs to be taken, patients were asked to complete a battery of self-report questionnaires about their back pain’s severity and their back disability. Pain severity was measured using two scales, one for pain during the entire time of the present episode and the other for usual pain during the previous week. In both cases, the pain scale was an 11-point numerical rating scale anchored at one end by the label “no pain” and at the other by “worst pain possible.” The question asked was, “We realize that there have been ‘good’ and ‘bad’ days, but, on average, how would you rate the severity of your back pain: (1) during the entire you have had it? (2) During the past week?” This format of questioning has been successfully used in previous studies and is a well-validated method of measuring a patient’s pain experience.

To measure back pain disability, patients completed the Revised Oswestry Disability Questionnaire (RODQ). The RODQ consists of 10 sections, each dealing with either pain or an area of daily activity likely to be affected by back pain. This is a well-validated measure.

**Radiographic Interpretation**

The lumbopelvic radiographs were interpreted by two qualified chiropractic radiologists, one at the Canadian Memorial Chiropractic College and the other at the Anglo-European College of Chiropractic. The radiologists were blinded as to the results on the questionnaires at the time of film interpretation. The inter- and intraexaminer reliability of determining the presence/absence of TLSV (an easy radiographic diagnosis), as well as the categorization of these anomalies, has already been studied using chiropractic radiologists and showed perfect reliability ($\kappa = 1$). Each set of radiographs was labeled as either “normal” or “TLSV” in terms of the presence/absence of this anomaly. If a TLSV was present, it was further categorized according to the method of Castellvi et al: type I, hyperplastic transverse process either unilaterally or bilaterally; type II, uni- or bilateral accessory joints; type III, uni or bilateral fusion; and type IV, accessory joint on one side and fusion on the other.
Table 1. Differences (student t test) between patients with and without TLSV in pain levels and individual questions on the RODQ

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>t</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain (entire)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TSLV+ (43)</td>
<td>5.767</td>
<td>2.287</td>
<td>0.59</td>
<td>.9532</td>
</tr>
<tr>
<td>TSLV − (308)</td>
<td>5.747</td>
<td>2.145</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain (past week)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TSLV+ (43)</td>
<td>5.744</td>
<td>2.441</td>
<td>−.388</td>
<td>.6985</td>
</tr>
<tr>
<td>TSLV − (308)</td>
<td>5.896</td>
<td>2.403</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Os 1 pain severity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TSLV+ (43)</td>
<td>2.977</td>
<td>1.318</td>
<td>−.489</td>
<td>.6255</td>
</tr>
<tr>
<td>TSLV − (309)</td>
<td>3.074</td>
<td>1.216</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Os 2 personal care</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TSLV+ (43)</td>
<td>1.419</td>
<td>1.239</td>
<td>−.001</td>
<td>.9991</td>
</tr>
<tr>
<td>TSLV − (308)</td>
<td>1.419</td>
<td>1.246</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Os 3 lifting</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TSLV+ (42)</td>
<td>2.143</td>
<td>1.719</td>
<td>−.313</td>
<td>.7545</td>
</tr>
<tr>
<td>TSLV − (307)</td>
<td>2.228</td>
<td>1.645</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Os 4 walking</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TSLV+ (43)</td>
<td>1.047</td>
<td>1.290</td>
<td>−.263</td>
<td>.7926</td>
</tr>
<tr>
<td>TSLV − (310)</td>
<td>1.100</td>
<td>1.244</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Os 5 sitting</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TSLV+ (42)</td>
<td>1.667</td>
<td>1.119</td>
<td>.211</td>
<td>.8331</td>
</tr>
<tr>
<td>TSLV − (310)</td>
<td>1.626</td>
<td>1.186</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Os 6 standing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TSLV+ (43)</td>
<td>1.953</td>
<td>1.272</td>
<td>.415</td>
<td>.6782</td>
</tr>
<tr>
<td>TSLV − (310)</td>
<td>2.048</td>
<td>1.421</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Os 7 sleeping</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TSLV+ (43)</td>
<td>1.558</td>
<td>1.259</td>
<td>1.100</td>
<td>.2719</td>
</tr>
<tr>
<td>TSLV − (310)</td>
<td>1.365</td>
<td>1.055</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Os 8 social life</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TSLV+ (43)</td>
<td>1.651</td>
<td>1.270</td>
<td>.231</td>
<td>.8175</td>
</tr>
<tr>
<td>TSLV − (308)</td>
<td>1.604</td>
<td>1.255</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Os 9 traveling</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TSLV+ (43)</td>
<td>1.372</td>
<td>.817</td>
<td>−.533</td>
<td>.5941</td>
</tr>
<tr>
<td>TSLV − (310)</td>
<td>1.455</td>
<td>.970</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Os 10 pain change</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TSLV+ (43)</td>
<td>2.698</td>
<td>1.166</td>
<td>−.506</td>
<td>.6128</td>
</tr>
<tr>
<td>TSLV − (306)</td>
<td>2.784</td>
<td>1.033</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Os total</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TSLV+ (41)</td>
<td>18.634</td>
<td>7.612</td>
<td>−.046</td>
<td>.9635</td>
</tr>
<tr>
<td>TSLV − (301)</td>
<td>18.691</td>
<td>7.445</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Values in parentheses are number of observations.
OS, Oswestry.

Statistical Analysis

Pain and disability levels were treated as interval data. Differences between patient groups in pain severity and disability were investigated using the unpaired t test. Multiple linear regression analyses were applied to investigate the effect of the presence of TLSV on pain and disability, controlling for the effects of age and sex.

RESULTS

Four hundred twenty-four patients were recruited to the study with 353 patients accepted after applying the exclusion criteria. The age range of the sample was 17 to 87 years (mean, 53.7 ± 16.27 years [SD]), and 51.0% of the sample were men. Forty-three patients had a transitional lumbosacral segment (12.2%) with 58% of these (25) being type II. There was no difference in age between the patients with and without TLSV (P = .5005) and no
difference in the ratio of the sexes \((P = .3408)\) between these two groups.

Using the Student \(t\) test to compare patients with LBP with and without TLSV, no statistical difference was found between these 2 groups for reported pain either from the previous week or for pain levels over the entire episode (Table 1). Similarly, there were no statistical differences between the 2 groups for any of the 10 subscales on the RODQ or for the overall (total) disability score (Table 1).

Controlling for age and sex, multiple regression analysis again showed no relationship between the presence of a TLSV and pain or disability on either of the pain scales or the RODQ. Age was a significant factor for both pain and disability, with more pain \((P = .039\) over entire episode) and disability \((P = .002)\) found in the older population. Sex had no effect on pain \((P = .761, P = .09)\) or disability \((P = .977)\) levels (Table 2).

**DISCUSSION**

The results of this study confirm that patients with transitional lumbosacral segments do not report more pain or disability than patients with normal lumbar morphology, despite that they have earlier and more severe degeneration at the motion segment above. These findings are consistent with those found by Luoma et al., who, while investigating male patients only, also had a large sample population. Although individual case reports and case series publications have noted that patients with TLSV often seem to have more pain particularly located on the same side as their anomalous articulation, these findings were not supported here. However, the specific side (if any) of a patient’s LBP was not recorded, nor was this compared with the specific side of anomalous articulation. Most patients with TLSV (58%) had a type II morphology, which included the presence of accessory joint/s with the sacral ala. Although Dai, reported that the incidence of type II transitional vertebrae was significantly higher in patients with low back, these findings were also not supported by the current study.

The results of this present study are not surprising when considering that previous work has failed to find a strong link between degenerative changes in the lumbar spine and pain and disability. Furthermore, the overall incidence of disk herniation is not higher in patients with a TLSV; it just has a strong predilection for the motion segment immediately above the anomaly and, thus, should not have contributed to any differences in pain or disability.

The incidence of transitional lumbosacral segments in this study (12.2%) is higher than that reported in the literature. This is because an adequate number of patients with this anomaly were needed to make statistical comparisons valid, and thus, data collection continued until this was accomplished. Comparison patients without the anomaly were easy to recruit, and data collection for this group ceased at an earlier date. Therefore, the 12.2% figure does not represent the true percentage within the population. Increasing the number of patients with this anomaly in this study should not have biased the results, as all radiographs were read blinded to any information from the questionnaires.

The main statistically significant results found in this study were for patient age and pain and disability levels. It is not at all surprising that when the multiple regression analysis was applied to determine the effects of age, sex, and the presence or absence of a TLSV, the older patients experienced significantly more pain and disability than younger patients. This was particularly true in the areas of total RODQ score; pain over the entire episode; the abilities of lifting, walking, and standing; as well as effect on their social life.

**CONCLUSIONS**

In this group of patients, there was no difference in pain or disability levels between those with and without transitional lumbosacral vertebra in spite of previous studies showing that similar patients have more degeneration at the motion segment immediately above this anomaly. Older patients in both categories generally report more pain and disability than younger patients.

**ACKNOWLEDGMENTS**

The authors thank the Anglo-European College of Chiropractic, Bournemouth, Dorset, England, and the Canadian Memorial Chiropractic College, Toronto, Ontario, Canada, for permitting us access to patients’ and radiographs.

**REFERENCES**

7. Winter AC. Categorization of lumbosacral transitional vertebrae and the incidence, location and severity of degenerative


10. Mikkelsen DG. Correlation between transitional lumbosacral vertebrae and degenerative changes at L4-5. Project for completion of BSc chiropractic degree. Bournemouth (UK): Anglo-European College of Chiropractic; 1996.


Erector Spinae and Quadratus Lumborum Muscle Endurance Tests and Supine Leg-Length Alignment Asymmetry: An Observational Study

Gary A. Knutson, DC,a and Edward Owens, DCb

ABSTRACT

Objective: To determine if there is an association between supine leg-length alignment (LLA) asymmetry and the endurance of the erector spinae (ES) and quadratus lumborum (QL) muscles.

Methods: Forty-seven subjects (21 women; average age, 36 years old) were tested for ES endurance using the Biering-Sorensen (B-S) test, and 69 (31 women; average age, 34.5 years) were tested for QL endurance. Subjects were examined for supine LLA and tested for ES and QL muscle endurance. The muscle endurance times were compared against those who did and did not demonstrate LLA asymmetry, and the side of the “short leg.”

Results: In the B-S test, volunteers with LLA asymmetry (n = 27) had a mean endurance time of 89.7 seconds (SD, 43.3), and the no-LLA asymmetry group (n = 20) had a mean endurance time of 161.5 seconds (SD, 57.1), a significant difference (P < .001). In the QL test, after correction for the effects of sex and exercise, those with a right “short leg” (n = 22) had a right QL endurance time of 25.9 seconds (SE, 4.2) and a left QL endurance time of 34.7 seconds (SE, 4.3). The right QL endurance time was significantly different from those subjects with balanced legs (P = .001). Those with a left “short leg” (n = 20) had a left QL endurance time of 28.6 seconds (SE, 4.7) and a right QL endurance time of 38.1 seconds (SE, 4.5). Both QL endurance times were significantly different from those with balanced leg-length (P = .002 and .016, respectively).

Conclusion: This study suggests that, using the B-S test, the group of volunteers who demonstrated a commonly used sign of subluxation/joint dysfunction, supine LLA asymmetry, had a decreased endurance times over those who did not. The QL endurance tests showed that the QL muscle ipsilateral to the supine short leg had significantly decreased endurance times over the same-side QL fatigue times in the no leg-length asymmetry group. (J Manipulative Physiol Ther 2005;28:575-581)

Key Indexing Terms: Muscles; Physical Endurance; Leg-length Inequality; Biering-Sorensen Test; Quadratus Lumborum

In 1984, a paper described a test, now called the Biering-Sorensen (B-S) test, which assesses the endurance of the erector spinae (ES) muscles.1 The B-S test involves positioning a subject prone on a table such that his or her upper body, above the level of the anterior superior iliac spine, is unsupported during the test. The pelvis, knees, and ankles are secured to the table with straps. The subject then holds the torso horizontal against the force of gravity, an average demand of 45% maximum voluntary contraction,2 for as long as possible. Biering-Sorensen found that those subjects with poorer isometric endurance on this trunk extension endurance test had a greater likelihood of developing low back trouble in the future.1 Further research showed the B-S test, in original and modified forms, to be both reliable3-7 and valid relative to predictive value.5,7 One study showed low reliability and noted that a weakness of the B-S test is the willingness of the subject to sustain the contraction; the authors commented that more highly motivated subjects “...are not necessarily those with the largest muscles or most efficient nervous systems.”8 However, Ebenbichler et al pointed to a series of studies that have been done investigating the surface electromyography (SEMG) fatigue pattern of the back muscles during submaximal voluntary contraction.9 SEMG changes
correlated with erector muscle fatigue, validating the subjective erector endurance tests against the objective SEMG. Moreau et al provide a literature review of the B-S test and write that the B-S test is cost-effective and easy to perform, can be done in little time, and requires no special equipment. They conclude that the B-S test is a useful measure for tracking changes in isometric extension endurance in the clinical setting.

Another test has been designed to evaluate the endurance of the quadratus lumborum (QL) muscle. The QL originates from the iliolumbar ligament, which itself is formed in infancy and childhood by a portion of the QL, the transverse processes of the lumbar vertebrae, and the 12th rib, and inserts on the posterior iliac crest. The action of the QL is to extend and laterally flex the spine, and, with the spine stable, to raise the pelvic crest. The QL has also been shown to be important for stabilizing the spine under load. Travell and Simons write that the QL “is one of the most commonly overlooked muscular sources of low back pain...”

The QL endurance test is performed by having the subject lie on his or her side supported by the arm, forearm at 90° to the torso. The upper leg is crossed in front of the lower, placing that foot in front of the lower leg foot. Using the arm as a brace and pivoting up at the ankles, the knee and pelvis are raised off the floor/table and held in a straight line as long as possible. This test causes a 54% maximal voluntary contraction of the QL ipsilateral to the down or support-arm side. McGill et al have found the QL fatigue test to be reliable, and the results have been found to be consistent with the endurance times of the more documented and validated B-S test.

One test or sign of spinal subluxation/joint dysfunction used by chiropractors is the leg check or alignment test. Leg check tests examine subjects for the so-called “short leg,” which is more aptly described as leg-length alignment (LLA) asymmetry in an unloaded posture. Supine physiological or functional LLA asymmetry is suspected of being caused by hypertonicity in suprapelvic and pelvic muscles, resulting in pelvic torsion or lateral flexion in an unloaded state. The suprapelvic muscles thought to be involved include the ES and the QL. Evidence points to intraexaminer and interexaminer reliabilities of the supine leg check, with validity relative to recurrent back pain, increased visual analog scale–rated intensity of back pain, and decreased SF-12 quality-of-life score. However, there is no consensus opinion.

Although both anatomical leg-length inequality (LLI) and LLA asymmetry are often referred to as a “short leg,” they seem to have different levels of incidence, physiological causes and structural consequences. Travell and Simons proposed that anatomical LLI leads to muscular overwork. Investigating this hypothesis, Mincer et al examined extensor muscle endurance on a group of subjects with x-ray–evidenced anatomical LLI. They found no difference in endurance times or neuromuscular control over a no LLI (≤2 mm) control group. These results raise questions as to whether overworked muscles are involved with small amounts of anatomical LLI. Given long-term loading, the anatomical short leg may be compensated for passively by changes in muscle length and bone/joint architecture via Heuter-Volkmann law and not actively via contractions of the lumbopelvic musculature.

Unlike the passive compensations for most anatomical LLI noted previously, the apparent LLA asymmetry in unloaded postures is proposed to be caused by unilateral muscular contraction. Tonic contraction of muscles, such as lifting the pelvic crest and causing the unloaded LLA asymmetry, would be expected to result in weakness/fatigue and pain. Evidence for the association of supine LLA asymmetry with increased intensity and recurrent pain has been reported. Asymmetry in fatigue characteristics and electromyography profiles of back muscles in LBP patients has also been demonstrated.

If supine LLA asymmetry is caused by suprapelvic muscle hypertonicity, one would expect to find reduced endurance of these muscles. Furthermore, if QL hypertonicity is responsible for lifting the hip and causing visual asymmetry in the unloaded supine LLA test, one would expect the QL muscle ipsilateral to the unloaded “short leg” to show reduced endurance. The purpose of this study is to test these hypotheses: supine LLA asymmetry is associated with decreased extensor endurance as evidenced by the B-S test and/or decreased endurance of the ipsilateral QL.

METHODS

Volunteers, recruited over a 4-month period at a private practice, were told that they would be performing moderate-level isometric exercise until fatigued. The exercises are considered safe, however, injury was possible (none was reported). The volunteers were thus informed, and consent was obtained. Exclusion criteria included acute antalgic low back pain, disk injury, radicular pain, and shoulder and ankle injury. Age, sex, and handedness of all volunteers were recorded. They were also asked about the presence of back and neck pain, the level of exercise they regularly engaged in, and whether their jobs involved physical activity.

The volunteers were not examined for the presence of an anatomical short leg for 2 reasons. The only reliably accurate way to do this is with specialized x-ray examination, and Rhodes et al demonstrated that the supine leg check was not correlated with radiographic anatomical LLI. As such, anatomical short leg, if present, would be random for both the LLA and no-LLA groups.

The supine LLA check was performed first. The volunteer stood at the end of the examination table, which is trapezoidal, wider at the foot (18 in) tapering to 8 in at the headpiece, standing 18 in off the floor, and is covered with
slick vinyl, and then sat down. Using their arms, they pulled themselves evenly toward the head of the table until just the ankles and feet extended off the table. The subject was then instructed to lie back so that the head rested on the headpiece and to relax. The procedure was demonstrated to those who were unfamiliar with it before their examination. With the subject supine, the examiner visually inspected the feet/legs for obvious asymmetry in y-axis translation (short leg) and foot rotations.

To perform the supine leg check, the examiner squats at the foot of the table, then lightly grasps and cups the heels of the subjects’ shoes. The feet are then derotated and squared to remove any foot rotation asymmetry. The examiner compares the positioning of the subject’s heel/sole interface from side to side. LLA asymmetry was then estimated, with less than 1/8 in considered to be even as was the protocol in the study of Hinson and Brown.22 The volunteers were selected to continue based on the presence of demonstrable LLA asymmetry; subjects with ambiguous findings were eliminated from the study. The presence of supine LLA asymmetry and what side appeared “short” were recorded, but the subject was not informed of those results.

The selected volunteers were then shown how the B-S test was performed. The subjects were positioned on a pelvic bench table with the torso above the ASIS unsupported except by the patient’s arms. The examiner held the subject’s ankles, and on signal, the subject raised their arms, placing them palm up on the table. Subjects were told to hold their torso horizontal to the point of excessive pain, fatigue, lack of ability to retain the horizontal position, or a maximum of 240 seconds, stopping the test by putting their arms back down for support.

During the rest period after the B-S test, the position and performance of the QL test were demonstrated. Subjects lay on their side with the ankle of the upper leg in front of the ankle of the lower leg. In the resting position, the torso is supported by arm (elbow bent at 90°). The superior arm crossed the chest to hold on to the supporting shoulder (deltoid). They were advised to get a firm and solid foundation from the supporting elbow to the shoulder. When positioned and ready, subjects raised the pelvis and knee off the table surface, pivoting at the shoulder and ankles to align the spine in a straight line. The subject was told to hold this position for as long as possible. The initial side for the QL test was determined randomly (coin flip, the next subject started on the opposite side). The subject performed the QL test until fatigued, rested for 2 minutes or more, and performed the test on the other side.

Because the examiner knew the results of the supine LLA test, verbal encouragement during the B-S and QL tests was avoided. Moreover, the subjects were not informed as to what was a “good” or “bad” time. The subjects were told only to hold the position for as long as possible.

Initial data analysis was performed in an Excel spreadsheet (Microsoft Corp, Redmond, Wash). The spreadsheet

Table 1. Biering-Sorensen data summary

<table>
<thead>
<tr>
<th>Variable</th>
<th>n</th>
<th>Mean</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>B-S holding time</td>
<td>47</td>
<td>120.2</td>
<td>61.3</td>
<td>32.0–240.0</td>
</tr>
<tr>
<td>B-S holding time</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>21</td>
<td>111.3</td>
<td>59.5</td>
<td>39.0–240.0</td>
</tr>
<tr>
<td>Men</td>
<td>26</td>
<td>127.5</td>
<td>61.8</td>
<td>32.0–240.0</td>
</tr>
<tr>
<td>Age</td>
<td>47</td>
<td>36.0</td>
<td>10.4</td>
<td>18–58</td>
</tr>
<tr>
<td>Categorical data</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LLA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left short</td>
<td>11</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Balanced</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right short</td>
<td>16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Handedness</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left</td>
<td>4</td>
<td>43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right</td>
<td>43</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exercise</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>10</td>
<td>28</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Moderate</td>
<td>28</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strenuous</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None or student</td>
<td>10</td>
<td>19</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>White collar</td>
<td>19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blue collar</td>
<td>18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>11</td>
<td>9</td>
<td>17</td>
<td>10</td>
</tr>
<tr>
<td>Neck pain</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low back pain</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neck and back pain</td>
<td>17</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Mean and adjusted B-S endurance times when supine LLA asymmetry is present or not

<table>
<thead>
<tr>
<th>B-S endurance times (s)</th>
<th>Mean (SD)</th>
<th>Adjusted mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>No LLA asymmetry</td>
<td>161.45 (57.7)</td>
<td>154.3a</td>
</tr>
<tr>
<td>(12 men, 7 women)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LLA asymmetry</td>
<td>89.7 (45.6)</td>
<td>95.0a</td>
</tr>
<tr>
<td>(14 men, 14 women)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Endurance times were adjusted for the effects of exercise and pain. 

* Significant difference (F2,43 = 13.9, P = .001).
RESULTS

B-S test data were collected on 47 subjects, 21 men and 26 women; age ranged from 18 to 58 years (mean, 36). The average endurance time for the B-S test was 120.2 seconds (SD, ±61.3). In this population, 27 had LLA asymmetry, and 20 did not. The mean values of the B-S test results and the breakdown of the categorical data are shown in Table 1.

The primary aim of data analysis was to detect any effect of LLA on the dependent variable of B-S holding time. Because other independent variables, such as age, sex, exercise, and job activities might also be related to back muscle endurance, Student t test and 1-way analysis of variance (ANOVA) were used to assess those effects. Significant interactions were found with exercise (F2,44 = 4.102, P = .023) and pain (F3,43 = 3.857, P = .016), but not with age, sex, job, or handedness. Analysis of covariance (ANCOVA) was used to test the effects of LLA on B-S, while adjusting for the relative effects of exercise and pain. ANCOVA omnibus test indicated that there was a significant difference between the leg-length groups (F2,44 = 13.909, P = .001). Patients with LLA exhibited significantly lower B-S holding times than those with balanced legs on the supine leg-length assessment (Table 2, Fig 1).

For the QL endurance tests, 69 subjects were examined. There were 38 men and 31 women, ranging in age from 18 to 58 years (mean, 34.5). QL holding time and other variables are described in Table 3.

The goal of the statistical analysis was to determine the relationship between LLA and QL endurance times. Other factors could also impact muscle endurance, including sex and level of fitness. As previously mentioned, t tests were used to examine the effects of the dichotomous variables, sex, and handedness. ANOVA was used to test the effects of the other demographic factors. Although handedness was not found to be related to QL fatigue times, sex had a statistically significant effect. QL fatigue times for men were nearly double those of women (Table 4). Of the other variables, exercise had a significant effect, whereas the subject’s age, the presence of back pain, and job activities did not.

Based on this initial analysis, ANCOVA was used to detect the effect of LLA on the QL times, with correction for the effects of sex and exercise. For the QL-left variable, the ANCOVA omnibus test indicated that there was a significant difference among the leg-length groups (F2,64 = 7.59, P = .001). Similarly, ANCOVA indicated a significant difference between the leg-length groups for the QL-right variable (F2,4 = 7.97, P = .001). A series of post hoc comparisons was conducted using pairwise comparisons controlling for type I error using the Bonferroni adjustment for multiple comparisons (familywise adjustment α = .05/3 = .017).

The results show that QL times are significantly lower when LLA is present, especially on the side ipsilateral to the short leg (Table 5). For the right QL, there is a significant difference between QL fatigue times only for the right short leg. In other words, when the subject’s right leg appears short, it significantly impacts the QL holding time on that same side, compared with subjects with neutral LLA, but not on the contralateral side. However, for the left QL, there is a significant difference when either leg is short. The effect is more pronounced on the ipsilateral side, that is, the difference between the left QL fatigue time when there is a short left leg is 23.0 seconds, whereas the difference is

### Table 3. QL data summary

<table>
<thead>
<tr>
<th>Variable</th>
<th>Continuous data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
</tr>
<tr>
<td>QL-left holding time</td>
<td>69</td>
</tr>
<tr>
<td>QL-right holding time</td>
<td>69</td>
</tr>
<tr>
<td>QL-left holding time</td>
<td>38</td>
</tr>
<tr>
<td>QL-right holding time</td>
<td>38</td>
</tr>
<tr>
<td>QL-left holding time</td>
<td>31</td>
</tr>
<tr>
<td>QL-right holding time</td>
<td>31</td>
</tr>
<tr>
<td>Age</td>
<td>69</td>
</tr>
</tbody>
</table>

### Categorical data

<table>
<thead>
<tr>
<th>LLA</th>
<th>Balanced</th>
<th>Right short</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left short</td>
<td>20</td>
<td>22</td>
</tr>
<tr>
<td>Handedness</td>
<td>Left</td>
<td>Right</td>
</tr>
<tr>
<td>Left</td>
<td>6</td>
<td>63</td>
</tr>
<tr>
<td>Exercise</td>
<td>None</td>
<td>Moderate</td>
</tr>
<tr>
<td>None</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>Job</td>
<td>None or student</td>
<td>White collar</td>
</tr>
<tr>
<td>None or student</td>
<td>18</td>
<td>26</td>
</tr>
<tr>
<td>Pain</td>
<td>None</td>
<td>Low back</td>
</tr>
<tr>
<td>None</td>
<td>15</td>
<td>28</td>
</tr>
<tr>
<td>Pain</td>
<td>Neck pain</td>
<td></td>
</tr>
<tr>
<td>Back pain</td>
<td>15</td>
<td>28</td>
</tr>
</tbody>
</table>

### Table 4. Relationship between sex and QL endurance times

<table>
<thead>
<tr>
<th>QL endurance times (s)</th>
<th>Men, n = 38 (SD)</th>
<th>Women, n = 31 (SD)</th>
<th>t(df)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Right QL</td>
<td>47.8 (28.1)</td>
<td>26.6 (14.6)</td>
<td>t _ L7 = −3.80</td>
<td>&lt;.001*</td>
</tr>
<tr>
<td>Left QL</td>
<td>50.5 (28.6)</td>
<td>26.1 (16.4)</td>
<td>t _ L7 = −4.21</td>
<td>&lt;.001*</td>
</tr>
</tbody>
</table>

* Significant.
only 17.0 for a right short leg (Fig 2). The columns of Table 5 containing the percentage reduction of QL times illustrate the reciprocal relationship. The mean QL fatigue times are reduced to 53% to 55% of the balanced QL on the side of the short leg but are 67% to 78% on the contralateral side.

**DISCUSSION**

The findings of this study show that, using the B-S test, subjects with supine LLA asymmetry had a significant decrease in endurance times compared with the no-LLA asymmetry group. In addition, the side of supine LLA asymmetry (the "short leg" QL) correlates more strongly with significantly decreased endurance of the ipsilateral QL. The subjects undergoing the endurance tests were blinded to the presence and side of LLA asymmetry, but the examiner was not. To overcome the lack of double blinding, the examiner refrained from coaching/encouraging the subjects during the endurance tests. Regardless, the potential for unintentional bias taints the results. Blinding the examiner of the muscle endurance tests to the results of the supine leg check would improve the significance of these findings and is recommended. Moreover, volunteers were selected to continue with the muscle function examination based on a demonstrable supine leg-length asymmetry check. Had all volunteers been examined, regardless of ambiguous leg check findings, the data may not have shown statistically significant relationships between the supine LLA check and muscle function tests.

Although age, sex, exercise frequency, or job type were examined for their effects on fatigue, the subject sample in this study was not selected with regard to balancing these variables between the LLA and no-LLA groups. Using a balanced sample would improve the findings. Still, the average endurance time in the B-S test for those without LLA asymmetry in this study (161.5 ± 57.1 seconds) was statistically the same (t test, \(P = .59\)) as the combined average in 14 other studies of normal subjects B-S test times (166.3 ± 41.5 seconds).\(^1\)

In the QL test, age, handedness or job activity did not have a significant effect on the endurance times. There is little research on the QL fatigue test compared with the B-S test; thus, any other potential confounders are not known but are likely to be similar to the B-S test. The confounders noted in the B-S test may apply bilaterally to the QL test (eg, exercisers would have longer fatigue times in both the right and left side QL). This was noted in 3 cases where the QL holding times were in longer than 100-second range, which is in contrast to the average of 40 seconds, but were increased bilaterally. In all 3 cases, the subjects were avid bicyclists. Unlike the B-S test, however, the QL test was used in this study as a comparison between right and left sides, noting asymmetry. Symmetry in back muscle function is considered normal.\(^3\) Symmetry in back muscle performance is considered normal, possible cofounders that may result in side-to-side asymmetry in QL fatigue times are more difficult to imagine. Possibly, unilateral sporting activities (golf, tennis, etc) and work duties could be confounding factors in side-to-side QL function asymmetry. An attempt was made to account for

| Table 5. Adjusted and percent difference values for the QL holding times as an effect of supine LLA asymmetry |
|----------------------------------|----------------------------------|
| QL tests                        | Adjusted right QL                | Adjusted left QL               |
|                                 | No. of subjects                  | Endurance time (SE)            | Percent of “balanced” | Endurance time (SE)            | Percent of “balanced” |
| Left LLA asymmetry              | 20                               | 38.1 (4.57)                    | 78.4                  | 28.6 (4.72)\(^a\)              | 55.3                  |
| Balanced LLA                    | 27                               | 48.6 (3.87)                    | 100                   | 51.7 (3.99)                     | 100                   |
| Right LLA asymmetry             | 22                               | 25.9 (4.20)\(^a\)             | 53.3                  | 34.7 (4.33)\(^a\)              | 67.1                  |

QL holding times were adjusted using ANCOVA for the effects of sex and exercise.

\(^a\) Significantly different from LLA = 0, \(P < .017\) (Bonferroni adjustment for multiple comparisons).

Fig 2. This figure shows the relationship between QL holding times and functional short leg (LLA). QL holding times are higher when there is no LLA present. There is a side-to-side relationship as well in that right QL times are lower than left when LLA is observed on the right; reciprocally, there is a decreased left QL time in the presence of LLA on the left.
greater strength of the dominant arm/side by recording hand use preference. Six of the 69 participants were left-handers; however, ANOVA found hand/side preference not to be related to QL fatigue times. The dominant side does not appear to be a confounding factor in this population. In future studies of QL fatigue times, screening for asymmetric sporting, work, and other activities should be included.

Stoppage of the tests because of pain other than in the back muscles occurred occasionally. In the QL test, pain in the supporting shoulder was the most common non–back-pain–related complaint. However, shoulder injury was an exclusion criteria, and there was an equal distribution of left and right LLA asymmetry, thus, a weak supporting shoulder would be random to the side of any LLA asymmetry.

QL endurance times were consistently and significantly shorter ipsilateral to the side of LLA asymmetry or supine “short leg.” In those subjects with right leg-length asymmetry, the left QL endurance time was also significantly reduced compared with the balanced leg-length controls. This was not the case in subjects with left LLA; the right QL fatigue time was not significantly reduced over the balanced leg-length group. We have no explanation for this finding.

Cooperstein and Lisi\(^\text{18}\) point out that many chiropractors and manipulation technique systems hold that the side of a posteriorly rotated ilium in the standing position identifies the side of the unloaded functional short leg, a view also held by some physical therapists.\(^\text{41}\) These opinions are likely based on the observation that inserting a lift under the foot elevates the acetabulum and causes posterior rotation of the ipsilateral, and anterior rotation of the contralateral hemipelvis.\(^\text{42}\) However, the forces acting on the pelvis from below (a lift) in a standing posture may not result in biomechanical changes similar to forces acting on the pelvis from above (QL muscle contraction) in an unloaded posture.

There appears to be greater QL fatigability (potentially indicating prolonged activity) ipsilateral to the supine LLA asymmetry (“short leg”). However, the action of the QL depends on whether the spine and/or pelvis is stabilized. In a loaded posture (standing), with the spine stable, a hypertonic QL would pull in a cephalad direction on the posterior aspect of the pelvic crest. This load would act to leverage the ipsilateral hemipelvis lower (known as an anterior superior ilium), causing the pelvis to torsion and having the opposite effect on the contralateral hemipelvis (a posterior inferior ilium). The degree of torsion while loaded, if any, would depend on the tension in the QL, the freedom of movement of the pelvis, and any preexisting pelvic torsion because of anatomical LLI. However, if this person now adopts an unloaded posture (supine or prone), the QL hypertonicity is free to lift or laterally flex the ipsilateral hemipelvis, producing LLA asymmetry or “short leg.” This idea is in agreement with that of Travell and Simons\(^\text{14}\) who write, “In recumbency, active TrPs [trigger points] shorten the [QL] muscle and can thus distort pelvic alignment, elevating the pelvis on the side of the tense muscle.”

Unloaded LLA asymmetry is believed to be caused by suprapelvic extensor muscle hypertonicity elevating the ipsilateral hip. Muscle hypertonicity can result in fatigue,\(^\text{30-32,35}\) which could explain the decreased endurance times of the ES and QL in this study. Muscle hypertonicity is also associated with pain.\(^\text{33,34}\) Mannion,\(^\text{43}\) in reviewing various causes of back pain writes, “…the evidence implicating highly fatigable back muscles in the development of low back pain is somewhat more substantial.” McGill\(^\text{31}\) notes that “…complete relaxation of the low back muscles is necessary to avoid compromising performance and an increased risk of musculoskeletal disorders.”

We propose that, in a subgroup of select subjects with nonradicular low back pain and decreased erector and/or asymmetric QL muscle endurance, the decreased endurance times are due to underlying hypertonicity. Long-standing ES hypertonicity could lead to pain and reduced performance on endurance tests; long-standing unilateral QL hypertonicity could lead to supine LLA asymmetry as well. The connection between LLA asymmetry and increased intensity of back pain has been noted previously,\(^\text{23}\) and the present study finds a connection between LLA asymmetry and endurance of the ES and QL.

Given the results of this study, a larger double-blind study of unloaded LLA asymmetry and muscle fatigue is recommended. An experimental design would be the next logical step, wherein B-S and QL times are measured before and after a course of chiropractic care. It would be particularly interesting to see if B-S and QL times increase or decrease, and QL times tend to balance out from side to side when the care has resulted in a prolonged restoration of functional leg-length balance. The data here suggest that this might happen but need confirmation from an experimental study.

**CONCLUSION**

In this study using the B-S erector muscle endurance test on a select population, a correlation was established between supine LLA asymmetry and a significant decrease in endurance times over the no-LLA asymmetry group. This study further showed that the side of the visual supine LLA asymmetry or “short leg” correlated most often with a decreased endurance performance of the ipsilateral QL. Further examination of muscle endurance relative to unloaded LLA asymmetry using a double-blind experimental methodology is recommended.

**REFERENCES**

EFFECT OF HIGH-INTENSITY STRENGTH-TRAINING ON FUNCTIONAL MEASURES OF BALANCE ABILITY IN BALANCE-IMPAIRED OLDER ADULTS

Jennifer A. Hess, DC, MPH, PhD, and Marjorie Woollacott, PhD

ABSTRACT

Objective: The aim of this study was to evaluate the effect of a 10-week, high-intensity strength-training program targeting key lower extremity muscles for the purpose of improving postural control in balance-impaired older adults.

Methods: A quasi-experimental, delayed entry controlled design was used to evaluate balance ability in balance-impaired older adults after participation in 10 weeks of high-intensity strength training focused on the quadriceps, hamstrings, tibialis anterior, and gastrocnemius muscles. Participants were evaluated using validated clinical measures of functional balance ability: the Berg Balance Scale, the Timed Up and Go, and the Activities-Specific Balance Confidence Scale.

Results: After strength training, the exercisers were significantly stronger than the control subjects. They improved significantly on the Berg Balance Scale (P = .030) from a mean score of 48.8 ± 2.4 of 56 before training to 51.2 ± 4.3 of 56 after training. The Timed Up and Go (P = .045) and the Activities-Specific Balance Confidence Scale (P = .038) also improved significantly in the experimental group. These changes are associated with a decrease in fall risk.

Conclusions: High-intensity strength training can safely and effectively strengthen lower extremity muscles in balance-impaired older adults, resulting in significant improvements in functional balance ability and decreased fall risk.

(J Manipulative Physiol Ther 2005;28:582-590)

Key Indexing Terms: Frail Elderly; Strength Training; Balance; Accidental Falls

Fifty percent of all people older than 80 years fall each year. In 2001, more than 1.6 million seniors were treated in emergency departments for fall-related injuries and falls are the leading cause of injury deaths. Of those who fall, 20% to 30% have moderate to severe injuries such as hip fractures or head injuries that decrease mobility and independence, and increase the risk for premature death. Because of these injuries, many individuals lose their independence. Fifty percent of fall injuries that require hospitalization result in discharge to a nursing home. Among people older than 75 years, those who fall are 4 to 5 times more likely to be admitted to a long-term care facility for a year or longer. These figures make it clear that clinicians are in need of rehabilitation tools that can prevent the occurrence of falls in balance-impaired older adults.

Diminished balance ability is multifactorial. It can be due to degeneration of visual and vestibular sensory systems, degeneration of proprioception or impairment of central processing or a combination of these factors. Yet, several studies have shown that lower extremity strength is a common factor associated with balance impairment in elderly fallers. For example, Lord et al found that ankle dorsiflexion strength was 1 of 3 variables that significantly discriminated between older adults who had no falls or one fall and those with a history of multiple falls. Pavol et al found that older adults who were classified as “after-step” fallers, in response to an induced trip, exhibited strength that is lower than average in lower extremity flexors and extensors. A study of nursing home residents with a history of falling found that muscle force (torque) and isokinetic power were significantly decreased in knee flexors (quadriceps [QD]) and extensors (hamstrings [HA]), and ankle
dorsiflexors (tibialis anterior [TA]) and plantar flexors (gastrocnemius [GA] and soleus). Dorsiflexors were particularly weak in fallers, suggesting that they are an important factor contributing to poor balance.12 Similarly, Wolfson et al13 evaluated falls in nursing home residents with a history of one or more falls in the last year and found that these individuals had less than half the ankle and knee strength of nonfallers and one tenth the ankle dorsiflexors strength of control subjects. These studies show a strong relationship between lower extremity strength and the ability to control posture.

The association between weak leg muscles and falling has led to a number of studies of strength training to enhance balance in balance-impaired older adults. These programs have evaluated the effects of strength training alone or in combination with other activities such as tai chi, aerobic exercise, and balance training.6,14-17 They have resulted in a variety of findings where some studies have failed to show any improvement in strength or balance control, even after 16 to 26 months of training, whereas others have shown improvements in both parameters. Many of these studies have had methodological limitations such as failing to strengthen key balance control muscles,15,18 providing an exercise program that benefited both subjects and control subjects, effectively diluting the training effect,6 unchallenging programs applied to minimally impaired subjects,19 or poor statistical power.20 Furthermore, a meta-analysis of several strength- and balance-training studies suggests that strength and endurance training might not affect the risk for falling at all, although balance training may.21

Fiatarone and Evans22 have suggested that some researchers, perhaps fearing injury to their subjects, have not used protocols of sufficient intensity to promote muscle strength gains. Hu and Woollacott,23 who studied multisensory balance training in older adults, suggest that training programs should target specific balance control systems and not use a shotgun approach. Postural control is defined as one’s ability to maintain the body’s center of gravity over the base of support during quiet stance and movement. Studies using movable platforms have shown that people respond to balance threats by activating set patterns of muscles and that older adults use different balance strategies than young adults.24-26 Young adults and healthy older adults rely more on hip muscles or strategies that combine the hip, ankle, and stepping.27,28 Although responses in ankle and knee flexors and extensors are critical to balance control, previous strength-training studies have not focused specifically on these muscles during strength-training paradigms. Based on these studies, exercise programs aimed at improving balance ability in balance-impaired older adults need to be of sufficient intensity to enhance strength and they must focus on key postural muscles in the lower extremities, especially ankle plantar and dorsiflexors.

The purpose of this study was to implement and clinically evaluate a high-intensity strength-training program for use as a clinical tool to improve balance control in balance-impaired older adults. It was hypothesized that 10 weeks of high-intensity strength training focused on key postural muscles in older adults with balance impairment would enhance clinical measures of postural control.

**METHODS**

**Subjects**

Participants were community-dwelling balance-impaired men and women, 74 to 96 years of age, recruited from the general older adult population in Eugene, Ore, by television, radio, and newspaper stories, and through talks at senior centers and health fairs. Participants were initially screened over the telephone to obtain pertinent medical history and obvious inclusion/exclusion information. A letter of approval to participate was sent to the primary care physician of all potential candidates before additional evaluation. Potential participants were then invited to receive neurologic and musculoskeletal examinations by a chiropractic physician.

Because of the complex nature of balance control, exclusion criteria were extensive. First, individuals who were unable to commit to a 10-week schedule of exercise 3 times a week were excluded. Health issues that could put subjects at risk for serious injury such as myocardial infarction in the last 6 months, recent angina, and poorly controlled hypertension (blood pressure [BP], >166/90 mm Hg) excluded participation. Factors that could confound research data including diabetic neuropathy, history of cerebral vascular accident (stroke), severe

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Experimental group (n = 13)</th>
<th>Control group (n = 14)</th>
<th>t Test (df = 25)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>3 M:10 F</td>
<td>2 M:12 F</td>
<td>0.569</td>
<td>.574</td>
</tr>
<tr>
<td>Age (y)</td>
<td>80.9 (5.8)</td>
<td>82.5 (6.9)</td>
<td>0.674</td>
<td>.506</td>
</tr>
<tr>
<td>Height (in)</td>
<td>64.1 (4.5)</td>
<td>63.2 (3.2)</td>
<td>-0.600</td>
<td>.554</td>
</tr>
<tr>
<td>Weight (lb)</td>
<td>151.2 (33.7)</td>
<td>149.3 (21.1)</td>
<td>-0.181</td>
<td>.858</td>
</tr>
<tr>
<td>BMI (kg/cm²)</td>
<td>25.5 (3.4)</td>
<td>26.2 (3.7)</td>
<td>0.495</td>
<td>.625</td>
</tr>
</tbody>
</table>

Values are presented as mean (SD) unless otherwise noted. M indicates male; F, female; BMI, body mass index.
congestive heart failure, Parkinson disease, diagnosed vestibular disorders, lower extremity joint replacements, or dementia were all reasons for exclusion. Potential participants were excluded if they used medicines known to impair balance ability, such as neuroleptics or benzodiazepines, or medications known to impair strength, such as prednisone. Participants were required to score a minimum of 25 of 30 on the Mini–Mental State Examination because mental competency was necessary to safely perform the exercise regimen. In addition, subjects needed to be ambulatory and were excluded if they were unable to walk 10 meters without assistance. Subjects were presented with a consent form before screening. This study was approved by the University of Oregon Committee for the Protection of Human Subjects Institutional Review Board. After the study was explained in detail, each participant signed a University of Oregon–approved 3-page consent form to ensure his or her willingness to participate in the study.

Clinical Measurement of Functional Balance Ability

Because balance impairment is difficult to quantify and may vary from person to person, participants were chosen based on having 2 of 5 selection criteria drawn from balance evaluation tools used in other studies. Participants were required to score: (1) 53 or less of 56 on the Berg balance scale and (2) 11.8 or more on the Timed Up and Go (TUG) test plus (3) a self-reported history of balance impairment (past falls; use of a cane, walker, or other walking aids; unsteady gait), (4) greater than 4 errors in a 2 meter tandem walk, or (5) gait abnormalities (one foot not clearing the other foot during gait, feet not clearing floor with gait, diminished arm swing, gait asymmetry). A fall was defined as any event that led to unplanned unexpected contact with the floor.

Practitioners rely on clinically obtainable measures to make decisions about functional balance ability in older adults. Several tools have shown utility for identifying individuals with functional motor performance impairment and were used in this study. The Berg Balance Scale is a validated clinical tool that consists of 14 functional subtests with a maximum score of 56 (excellent balance), which is used clinically to quantify balance impairment in the older adults. It requires timed performance of a variety of tasks, including sit-to-stand, reaching, leaning over, turning, and stepping. The Berg Balance Scale has been shown to have strong interrater and intrarater reliability ($r = 0.76$). Although it provides good specificity for predicting who will not fall, it has been shown to have poor sensitivity for predicting who would fall. Yet, combined with a self-reported history of imbalance, the Berg has been shown to be a sensitive and specific tool for predicting fall risk.

The TUG is another test for assessing physical mobility in geriatric populations. To perform the TUG, an individual is timed while rising from a chair, walking 3 meters, returning to the chair, and sitting down. The TUG has good interrater and intrarater reliability, and is a reliable and valid measure of functional mobility. It also is a sensitive and specific measure of faller status in community-dwelling balance-impaired older adults.

The Activities-Specific Balance Confidence Scale (ABC) is a balance perception questionnaire designed with input from clinicians and community seniors, which has shown good test-retest reliability ($r = 0.92$, $P < .001$). It contains 16 questions asking about self-confidence when performing various activities of daily living that require balance. Respondents are asked to rate their percent confidence where 1% is not confident and 100% is very confident. These 16 scores are averaged for a total score. Those scoring less than 68% are considered to have low mobility skills. The ABC is internally consistent and showed good test-retest reliability and validity for discriminating loss of balance confidence in more highly functioning older adults.

Study Design and Procedures

This study used a nonrandomized, quasi-experimental, delayed entry design. Initially, 4 individuals were enrolled as subjects in a pilot study. Two men and two women completed the 10-week, high-intensity strength-training program. After the pilot, individuals were enrolled in the larger study. It became apparent at the onset of recruitment that enrollment of participants solely as control subjects would be difficult. Most people refused to enter the study if they could not participate in strength training. Therefore, 7 study participants served as control subjects before serving as experimental subjects. The remaining 7 control subjects and 6 experimental subjects (4 of whom were pilot subjects) were different individuals. As such, this study had elements of both within-subject and between-subject designs. Delayed entry refers to the enrollment of 11 control subjects and 4 pilot experimental subjects in the spring of the year, followed by enrollment of the remaining 9 experimental and 3 control subjects in the autumn of the year. Delayed entry was necessary to maintain compliance. The strength-training regimen required a commitment of 3 visits per week for 10 weeks, in addition to data collection before and after this time.

Leg strength was measured in the experimental and control groups at the onset of the study and, again, 10 weeks later. Strength in the QDs and HMs, TA, and GA was quantified using a test of one repetition maximum (1RM) strength on the same equipment the experimental group would use for strength training. Control subjects were instructed not to begin or partake in any exercise programs during the 10 weeks they were control subjects. Three months after the completion of the control phase, individuals who served as control subjects became part of the experimental group. This break was included to minimize attrition and
STRENGTH-TRAINING PROTOCOL

Experimental group subjects participated in a 10-week, high-intensity strength-training program that consisted of specific exercises using professional equipment: (1) the Hammerstrength tibia dorsiflexion machine (Hammerstrength, Schiller Park, Ill), (2) Maxicam machine (Muscle Dynamics, Torrance, Calif) for strengthening plantar flexors (GA), and (3) Maxicam variable-resistance machines for performing knee extension to strengthen QDs and knee flexion to strengthen HMs. Each experimental subject was personally supervised during all exercise sessions by pre–physical therapy students trained in safe and proper exercise performance. The training protocol was modeled after a format developed by Fiatarone et al37,38, which has been shown to be safe and effective for training balance-impaired older adults and follows American College of Sports Medicine39 guidelines for strength training in seniors.

To prevent injury to subjects by having them perform a true 1RM (defined as the maximum weight a person can lift only one time before becoming fatigued), an estimated 1RM was used following a protocol designed by Lombardi.40 This method allows individuals to use multiple repetitions (8-10) at lower weight intensities and a table to predict the true 1RM. The Lombardi table is an exponential form using the reciprocal of the number of repetitions to the 0.1 power to estimate the percentage of 1RM. This method is not only safer but also an accurate method of determining the maximum weight an individual can lift.41 The 1RM concept has been safely used in other strength studies of older adults without incidence of injury.42-44

The week before training, experimental subjects were given a practice session on the training equipment. Subjects trained 3 times a week, performing 3 sets of 8 repetitions for each muscle group, over a 10-week period. In week 1, subjects exercised at 50% of their estimated 1RM to prepare them for more rigorous exercises and, as another measure, to avoid injury. Beginning in week 2, all subjects trained at 80% of their estimated 1RM. The training stimulus was increased each week or as tolerated, whereas a 1RM level was evaluated every 2 weeks to keep the load constant at 80% effort throughout the training period. For consistency, subjects were asked to perform each exercise repetition quickly, in approximately 6 seconds, and to perform repetitions in a consistent, evenly paced manner. For example, for ankle dorsiflexor training, they were instructed to raise the weight with the legs in 2 seconds then lower the weight during the eccentric phase of contraction in approximately 4 seconds. They were instructed to pause approximately 2 seconds between repetitions, and all exercises were performed bilaterally. A 2-minute rest was given between each set of 8 repetitions. Before exercising, subjects spent 20 minutes warming-up; 10 minutes on an ergometer bicycle, and 10 minutes of stretching the legs and back. In addition, because exertion can increase BP, especially toward the last repetitions, there was intermittent monitoring of pulse rate and BP. Each training session lasted approximately 1 hour.

Because of the increased nutritional demands associated with strength training, snacks of water, fruit juice, and nuts were supplied. Recommendations were made to both exercisers and control subjects to supplement their diets with a multivitamin, eat plenty of protein, and drink plenty of water, general recommendations applicable to all older adults for good health but especially important for those actively exercising.

Data Analysis

The predictor variable was a change in 1RM strength in response to 10 weeks of high-intensity strength training. Specifically, changes in the strength of the TA, GA, QDs, and HMs were quantified. Proportional changes in muscles strength were evaluated by subtracting prestrength from poststrength and dividing the outcome by prestrength. The proportional change in strength was used to compare strength changes across individuals because participants varied widely in their initial strength and total strength gain. The response variables were the Berg Balance Scale, the TUG, and the ABC scores.
Sample size was determined by reviewing studies on postural responses in balance-impaired older adults, which used similar exercise-training regimens and clinical outcome variables. SPSS 11.5 statistical software (Chicago, Ill) was used to determine differences in clinical balance tests between experimental and control subjects. A Mann-Whitney nonparametric test for two independent samples was used to evaluate training differences in the experimental and control groups for changes in strength and clinical measures. To obtain blinded results, the Berg balance score was also scored independently by a pre–physical therapy student from video of each participant.

**RESULTS**

This study population consisted of 13 experimental subjects and 14 control subjects. Seven individuals served in both the control and experimental groups, whereas an additional 6 experimental and 7 control subjects were separate individuals. Compliance in completing the study was generally good, although 4 people dropped out after being control subjects, two for health reasons and two for personal reasons. Ninety-three percent of each group, experimental and control, finished the study. In the experimental group, there was 100% training compliance. There were no injuries as a result of the high-intensity strength-training program. There were no statistically significant differences in age, height, weight, or body mass index between subjects in the experimental and control groups (Table 1).

Strength in ankle and knee flexors and extensors was not significantly different in experimental and control groups before strength training (Table 2). Before training, TA 1RM strength was $15.6 \pm 13.8$ lb in experimental and $13.7 \pm 3.7$ lb in control subjects, whereas GA 1RM strength was $45.4 \pm 15.5$ lb in experimental and $60.8 \pm 23.4$ lb in control subjects. For the thigh muscles, pretraining QD 1RM strength was $60.3 \pm 8.1$ lb in experimental and $74.4 \pm 27.1$ lb in control subjects, whereas HM 1RM was $37.1 \pm 12.9$ lb in experimental and $39.0 \pm 8.5$ lb in control subjects. The 1RM strength for TA and GA increased significantly in experimental compared with control group subjects. Tibialis anterior 1RM strength after training was $26.8 \pm 16.1$ lb in experimental and $14.3 \pm 2.3$ lb in control subjects, whereas GA 1RM strength was $94.2 \pm 24.6$ lb in experimental and only $51.7 \pm 9.3$ lb in the control group (Fig 1).

The 1RM strength also increased significantly in the QDs and HMs. Posttraining QD 1RM strength was $26.8 \pm 16.1$ lb in experimental and $14.3 \pm 2.3$ lb in control groups, whereas HM 1RM was $66.7 \pm 24.9$ lb in experimental and $39.9 \pm 9.5$ lb in control groups.

In the experimental group, mean GA 1RM strength change was significantly correlated with mean Berg Balance Scale scores (Pearson correlation coefficient = $-0.683$, $P = .014$; Kendall $\tau = -.605$, $P = .007$). There-
before, Berg scores increased with increasing GA strength. However, no other correlations were found between strength and clinical measures.

There were no significant differences in Berg balance scores between experimental and control groups before strength training (Table 2). The mean Berg score for the experimental group pretraining was 48.8 ± 2.4 of 56, whereas for control subjects, it was 48.5 ± 2.8 of 56. After 10 weeks of strength training, the experimental group scored significantly higher on the Berg Balance Scale compared with control subjects with a similar baseline (Fig 2A and B). After training, the experimental group averaged 51.2 ± 4.3 of 56 on the Berg, whereas control subjects averaged 49.5 ± 3.0 of 56.

Before strength training, there were no significant differences in TUG scores between experimental and control groups (Table 2). The experimental group mean TUG score before training was 11.5 ± 2.4 seconds, whereas for control subjects, it was 11.2 ± 1.7 seconds. When measured after 10 weeks of strength training, there were significant differences in the experimental group’s mean TUG scores (Fig 3) where their values decreased to an average of 9.7 ± 2.5 seconds, whereas control subjects scored 11.8 ± 3.3 seconds.

Study participants rated their self-perception of balance ability during activities of daily living using the ABC scale. There was no significant difference in ABC scores between the experimental and control groups before the strength-training program (Table 2). The experimental group scored an average of 80.3% ± 15.0% before training, whereas control subjects scored on average 81.1% ± 11.7%. After completing the strength-training program, the experimental group rated themselves significantly higher in balance self-confidence (Fig 4). Their mean ABC score after training was 88.3% ± 10.3%, whereas control subjects were unchanged with a mean score of 81.2% ± 13.5%.

**DISCUSSION**

**Strength**

Strength declines approximately 15% per decade between the ages of 50 and 70 years, and approximately 1.5% per year after the age of 70 years, and decreases in strength are associated with falling in elderly people. Örlander and Aniansson concluded that muscle remains trainable regardless of age, although some studies have failed to find gains in strength with training. This study evaluated the effect of a 10-week, high-intensity strength-training program designed to enhance functional balance control in balance-impaired older adults. Four key posture control muscles were targeted, the TA, GA, QD, and HM. This program was unique because, in addition to being an aggressive program with balance-impaired participants, ankle musculature was emphasized. Subjects used the Hammerstrength Tibialis Anterior Machine (Hammerstrength) that isolated the TA for maximum strength gain.

The training regimen resulted in statistically significant increases in strength in ankle and knee flexors and extensors in this population of balance-impaired individuals. Of particular interest was ankle flexor and extensor strength because these muscles are linked to balance impairment in older adults with a history of falling. The TA as an ankle dorsiflexor is important for balance stabilization during backward slips in which anterior ankle strength is needed to prevent the body’s center of mass from moving posteriorly beyond the base of support. Tibialis anterior strength increased 42% in subjects after 10 weeks of training. Conversely, the GA is an ankle plantarflexor that concentrically contracts during forward trips to prevent the center of mass from moving anteriorly beyond the base of support. After 10 weeks of training, GA strength improved 52%. Similarly, strength in the QDs, knee extensors that are important for controlling backward sway, improved 35%, whereas in the HMs, knee flexors that attenuate forward sway, strength increased 44%.

Other studies have shown that older adults can safely perform high-intensity strength training, resulting in muscle strength gains, but findings are mixed as to the effect of strength training on balance control. Judge et al found that resistance training enhanced strength but it did not affect balance as measured by gait velocity or chair rise time, but the population evaluated was not balance-impaired. Fiatarone and Evans suggest that some training programs may not be of sufficient intensity to promote strength gains. It is also possible that exercise equipment does not target key postural control muscles such as the TA. Indeed, most commercial gyms and rehabilitation centers do not have equipment that specifically works the TA. For example, Buchner et al used resistance strength training of moderate intensity (2 sets of 10 repetitions at 60%-75% 1RM for 6 months) in older adults for the purpose of enhancing balance control. They did not find that strength training enhanced balance...
ability, but they also did not find significant increases in isokinetic ankle dorsiflexor strength after 6 months of training. Other studies of exercise programs have focused on aerobic capacity or endurance rather than maximizing strength, which may be a reason that no improvement in balance ability was noted. In this study, strength-training equipment was used that specifically targeted muscles known to be weak in people with a history of falling. The program pace and intensity were such that subjects showed substantial gains in strength over a few weeks.

A variety of mechanisms may be responsible for improvements in functional balance control in these older adults. Enoka suggests that training-related improvements in muscle strength may be due to changes in neural mechanisms such as enhanced output from supraspinal centers, altered neural drive that decreases coactivation of antagonist muscles, and greater activation of synergist muscles, or more effective coupling within the spinal interneuronal pathways, producing greater cross-limb education. Sale asserts that increased neural activation of prime movers increases motor unit synchronization and enhances activation of fast-twitch motor units, resulting in better muscle coordination and efficiency. Similarly, Connelly and Vandervoort conclude that changes in ankle muscle force are due to adaptation of neural factors such as changes in recruitment patterns of activated motor units and enhanced synchronization of motor unit firing.

Conversely, others believe changes in balance ability in older adults in response to gains in strength are due to higher rates of torque development possibly caused by larger type II/I fiber ratios or greater cross-sectional area of type II muscle fibers. Thelen et al. found that maximum rates of ankle torque generation were 25% to 36% lower in older adults compared with younger adults. Strength training has been shown to increase the magnitude and rate of force production in lower extremity muscles. Furthermore, it has been shown that balance recovery limits increase as ankle torque and the rate of ankle torque production increase. Although it was beyond the scope of this study to isolate the mechanism responsible for the strength gains found in these older adults, it seems likely that a variety of neural and biomechanical adaptations contributed to the improved functional balance control noted in this population.

Functional Clinical Tests of Balance

Two clinical tests routinely used to gauge functional balance ability, the Berg and the TUG, both improved significantly in subjects who underwent 10 weeks of high-intensity strength training. In those scoring between 43 and 52 on the Berg Balance Scale, each one-point decrease in score has been equated with a 2% to 7% increase in fall risk. A one- to two-point increase is substantial when one considers the ramifications of fall injuries among older adults. Subjects who strengthened their QD, HM, TA, and GA, key postural muscles, showed an average of 2.4 points increase in their Berg scores, which translates into a 4.8% to 16.8% decrease in fall risk in 10 weeks, in response to a 43% overall increase in muscle strength.

The TUG test is both sensitive and specific for identifying individuals with balance impairment who are likely to fall. Shumway-Cook et al. showed that older adults with scores of greater than 13.5 seconds have a 90% probability of being fallers. This study did not specifically target fallers; rather, community-dwelling older adults with moderate balance impairments were the focus. In the experimental group, a mean pretraining TUG score of 11.5 seconds corresponds to a fall risk of approximately 50%. After strength training, the mean TUG score in the experimental group was 9.7 seconds, which corresponds to approximately a 20% fall risk. Therefore, as a result of strength training, using the TUG as a measure of fall risk, those in the experimental group decreased their fall risk by approximately 30%.

The third clinical measure evaluated in this study was a self-perception survey of balance confidence, the ABC. In conjunction with the Berg test, the ABC has been shown to be highly sensitive (89%) and specific (96%) for predicting fall. A person with an ABC score of 87.5% or more is considered to be a healthy older adult. In this study, in response to high-intensity strength training, the mean ABC score in the experimental group increased significantly from 80.3% before training to 88.3% after strength training, well higher than the confidence level for a healthy older adult. Although this study did not measure the effect of strength training on actual numbers of falls, the results of these 3 clinical tests do suggest that better lower extremity strength results in increased functional balance control ability for this group of moderately balance-impaired adults.

The level of initial balance impairment was not quantified as part of this study; only the change in balance was determined. To be considered for this study, participants had to meet certain criteria frequently used to assess balance impairment among community-dwelling adults. Yet, during the strength-training program, differences in balance impairment between subjects became obvious. The mean age of subjects was 81 years. The oldest and most balance-impaired subject was a 94-year-old woman, compared with the youngest subject who was 75 years of age. While the more balance-impaired individuals performed well, their gains over the 10 weeks were less substantial than gains made by the more vigorous individuals. Hence, a longer training program for the more balance-impaired individuals may have resulted in greater strength gains for those people. The choice of a 10-week strength-training program was based on training durations used in other studies that resulted in significant improvements in strength and due to practical considerations such as the ability to maintain compliance. There is nothing in this study to suggest that subjects reached a plateau in strength gains, and this is particularly true for the more balance-impaired individuals.
who, given more time, may have achieved greater gains, both in strength and in balance control.

Clinical Importance

The findings from this study are important for clinicians for 3 reasons. First, they draw attention to clinical tests, the BERG, TUG, and ABC, which are readily usable in a doctor’s office to assess a patient’s functional balance ability. Second, many gyms and fitness centers are beginning to cater to older populations, yet some clinicians have been hesitant to advise high-intensity strength training for fear of adverse health effects or injury. This study highlights that older adults are able to safely perform vigorous exercise and strength training. Third, and most important, high-intensity strength training provides clinicians with a tool to conservatively manage balance impairment in older patient populations. Based on the results reported here, strength training can reduce fall risk by 5% to 20% in only 10 weeks while enhancing activity confidence to healthy levels.

Study Limitations

There was also no correlation established between improvement in the functional parameters evaluated in these subjects and actual fall rates. A prospective study that follows participants for a period after training to quantify fall rate would strengthen the generalizability of the results to fall prevention. However, that type of follow-up was beyond the scope of this study and would depend upon compliance with an ongoing strength-training program.

Conclusion

Thirteen balance-impaired older adults safely completed a high-intensity strength-training program and significantly enhanced their lower extremity strength. These strength gains resulted in significant improvements on 3 clinical measures of functional balance ability. High-intensity strength training may provide clinicians with a viable option for improving functional balance control that will help their patients remain independent, active, and healthy.

References

35. Shumway-Cook A, Brauer S, Woollacott MH. Predicting the probability for falls in community dwelling older adults using the up and go test. Phys Ther 2000;80:896-903.
37. Fiatarone MA, Marks EC, Ryan ND, Meredith CN, Lipsitz LA, Evans WJ. High-intensity strength-training in nonagenarians. JAMA 1990;263:3029-34.
Measurement of the Pressure Applied During Motion Palpation and Reliability for Cervical Spine Rotation

Justin Marcotte, DC, a Martin C. Normand, DC, PhD, b and Pierre Black, MSc c

Abstract

Objective: To measure the pressure applied during motion palpation for cervical spine rotation and to verify its effect on reliability when the kinematics of the test are standardized.

Methods: The pressure of palpation used during the test of cervical spine rotation was measured by means of flexible and extra-fine pressure sensors linked to an electronic interface. Seven pressure measurements (left rotation from C1 to C7) for each of 24 examiners were taken. In addition, the examiners were asked to detect the presence of intervertebral fixation while palpating.

Results: Pressure of palpation varied from 4.0 to 41.0 N/cm² among the examiners. Standardization of the kinematics of the test establishes a strong reliability of identifying a fixation (κ varying from 0.701 to 0.748).

Conclusions: The pressure applied during motion palpation for cervical spine rotation is light to moderate. It can vary tenfold (4-41 N/cm²) and remain reliable for identifying a fixation as long as the kinematics of the test are standardized.

(J Manipulative Physiol Ther 2005;28:591-596)

Key Indexing Terms: Reproducibility of Results; Cervical Vertebrae; Force; Pressure Sensors

At the Université du Québec à Trois-Rivières (UQTR), emphasis is placed on the technique of motion palpation to establish the criteria for chiropractic adjustment of the manipulable lesion. To our knowledge, accredited chiropractic teaching institutions include in their academic curriculum various examining techniques of palpation to test for functional lesions of variable spinal areas. The goal of this study is to measure the pressure of palpation applied during motion palpation for rotation of the cervical spine and to verify its effect on the reliability of the test.

In reliability studies, motion palpation has been fraught with difficulties related to the standardization of the procedure and to the subtlety of vertebral lesions. Thus, Mior et al,1,2 Mootz et al,3 and Nansel et al4 showed the weak or clinically insignificant reliability of the test for certain groups of young, asymptomatic subjects. Herzog et al5 and Deboer et al6 obtained mixed results. However, other researchers such as Wiles7 and Love and Brodeur8 verified an acceptable reliability for their own studies of motion palpation testing.

The limited success of these studies to verify the experimental reliability of spinal motion palpation casts doubt on the capacity of examiners to reproduce various technical parameters of the test. O’Malley9 stipulated that a questionable aspect of the vertebral subluxation could arise from the theoretical preconceptions of the examiners. Cassidy and Potter10 are of the opinion that the subtle nature of the subluxation and the technical difficulties inherent in the testing for it have largely complicated a favorable experimental methodology. Panzer11 attributed the stronger intra-examiner reliability to a technical or interpretative skew on behalf of the examiner. Panzer11 and Dishman12 invited future researchers to standardize the technical parameters of the test before subjecting it to an interexaminer reliability study.

Our previous studies13,14 showed that a higher level of standardization of the kinematics of the test and the use of subjects with a history of mechanical disorders of the area

---

a Professor, Département de chiropratique, Université du Québec à Trois-Rivières, Quebec, Canada.
b Professor, Département des Sciences de l’activité physique, Université du Québec à Trois-Rivières, Quebec, Canada.
c Assistant Professor, Département des Sciences de l’activité physique, Université du Québec à Trois-Rivières, Quebec, Canada.
Sources of support: This study was made possible by a grant from the Fondation Chiropratique du Québec.
Submit requests for reprints to: Dr. Justin Marcotte, Département de chiropratique, Université du Québec à Trois-Rivières, CP 500, Trois-Rivières, Québec, Canada G9A 5H7 (e-mail: justin_marcotte@videotron.ca).
0161-4754/$30.00
Copyright © 2005 by National University of Health Sciences.
doi:10.1016/j.jmpt.2005.08.003
tested significantly improved the reliability of the test. In these experiments, the examiners underwent supervised technical training until a success of reproducibility, that is, standardization, was achieved. In our studies on standardization of cervical spine rotation, the plane of palpation was proved to deviate by no more than 6° of inclination from the transverse plane of rotation (as defined by the international system of coordinates), when standardization and reliability were achieved.

Another important variable to consider when testing for motion palpation is the pressure of palpation applied when the examiner contacts the vertebra and evaluates the end feel at each intervertebral level. To our knowledge, this component of the test has not been the subject of previous research. This study evaluates, by instrumented measurement, the pressure applied during the test of motion palpation for cervical spine rotation. At the same time, the purpose of the study was to verify the effect of the pressure of palpation on the interexaminer reliability when standardization is achieved.

METHODS

Subject Selection

The participants in the study were selected from the students and professors of the Doctor of Chiropractic Program at UQTR. Informed consent was obtained from all the participants. The research project was assessed and approved by the institutional review board of UQTR and conforms to the Declaration of Helsinki.

Twenty-four examiners were selected to perform 7 pressure evaluations each. Twenty-three student examiners, in addition to an experienced chiropractor, were accepted after having successfully completed a maximum of 12 hours of supervised training to comply with the criteria of standardization as set forth in our previous study.14 The examiner, stationed at the head of the supine subject, conducted rotatory motion palpation of the cervical spine. The procedure was monitored by camera and markers allowed measurement of the kinematics. The criterion of acceptability for the examiners was their successful reproducibility of the kinematics of the test, that is, the plane of palpation had to be maintained within 6° of the transverse plane of rotation.

Other Participants

Three asymptomatic male patients (22, 25, and 28 years of age and of similar height and weight) were selected after an examination by the experienced chiropractor. They all had a history of occasional neck pain of mechanical origin and presented notable vertebral fixations in the plane of left cervical spine rotation. Each patient was examined by only a third of the examiners to minimize the potential discomfort of repeated examination.

Equipment

Pressure sensors (Interlink Electronics, Camarillo, Calif) twinned with an electronic interface were used for this experiment. The sensors have a surface area of 0.20 cm². These simulate the lateral palmar surface of the distal interphalangeal joint of the index finger (the contact point), which was considered the useful surface for palpation in our experiment. The sensors are extra thin, 0.20 mm in thickness (0.008 in), and flexible to minimize interference for the contacting finger. They are sensitive to forces from 0.2 to 100 N (weight of 20 g to 10 kg) and to pressures from 1 to 100 N/cm² (1.5-150 psi). The margin of error for the system was 2% to 5%, excellent for data to be collected in this experiment.
Experimental Design

The test of motion palpation in cervical spine rotation was executed in the usual way with the subject lying supine in a passive attitude. The examiners were asked to use a force that enabled them to adequately feel their palpation yet remained comfortable for the patient. The point of contact used by the examiners (the lateropalmar surface of the distal interphalangeal joint of the right index finger) had a pressure sensor attached and connected to a computerized system of data collection (Figs 1 and 2).

The examiners had to note, blinded to the other examiners, the intervertebral levels presenting with vertebral fixation. Mild or dubious fixations were to be considered as normal and were not retained.

Our strategy for accurate identification of the level of fixation was influenced by 2 previous studies. The identification of the articular pillars by marking of the skin proved to be an unfruitful practice because of the cutaneous movement in reference to the underlying osseous structures during motion palpation. The grouping by area approach is too arbitrary and would inflate the reliability for certain areas. For these reasons, we accepted a margin of examiner uncertainty for positive findings of segment. Because of the presence of easily identifiable osseous landmarks, fixations at C1 and C7 had to correlate precisely.

Data Analysis

The pressure of palpation results was compiled in table and graph form. Descriptive statistics were used in regard to the analysis of these data. The data for reliability study were analyzed using the coefficient of concordance \( \kappa \). Matrices of compilation were used, where the values \( Q \) and \( T \) represent the number of agreements between examiners; \( R \) and \( S \) represent the number of interexaminer disagreements. Reliability results from calculations of the percent agreement and the coefficient of concordance \( \kappa \) are indicated in the corresponding tables. The power of the prediction (\( P \) value) for the data collected was \( P < .01 \). The clinical relevance of the coefficient of concordance was offered by the following values of \( \kappa \): \( \kappa < 0.0 = \text{null}, 0.0 < \kappa < 0.2 = \text{weak}, 0.2 < \kappa < 0.4 = \text{fair}, 0.4 < \kappa < 0.6 = \text{moderate}, 0.6 < \kappa < 0.8 = \text{strong}, 0.8 < \kappa < 1.0 = \text{very strong} \).

RESULTS

Pressure of Palpation

The surface area of palpation corresponding to the point of contact on the posterior aspect of the articular pillar of the contacted vertebra was estimated at 0.2 cm\(^2\), which corresponds to the surface area of the sensors. Larger sensors would be indicated for larger contact points (lumbar region or sacroiliacs). The pressure of palpation recorded by our measuring instrument is expressed in newton per square centimeter.

The results revealed that, uniformly, the examiners begin their palpation with a minimal pressure that rises smoothly and peaks when checking for joint play (end feel). We insured that the palpation remained comfortable for the patient by asking them their impressions after palpation. Fig 3 provides a visual representation of the pressure of palpation applied in a typical series of tests.

Table 1. Pressure of palpation (N/cm\(^2\))

<table>
<thead>
<tr>
<th>Series</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>C4</th>
<th>C5</th>
<th>C6</th>
<th>C7</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20.0</td>
<td>21.0</td>
<td>14.0</td>
<td>20.5</td>
<td>21.0</td>
<td>20.0</td>
<td>15.0</td>
</tr>
<tr>
<td>2</td>
<td>15.5</td>
<td>17.0</td>
<td>19.0</td>
<td>27.0</td>
<td>21.5</td>
<td>13.5</td>
<td>17.5</td>
</tr>
<tr>
<td>3</td>
<td>24.5</td>
<td>32.5</td>
<td>39.5</td>
<td>20.5</td>
<td>24.0</td>
<td>11.0</td>
<td>21.5</td>
</tr>
<tr>
<td>4</td>
<td>21.0</td>
<td>19.5</td>
<td>22.5</td>
<td>26.0</td>
<td>14.5</td>
<td>12.0</td>
<td>19.5</td>
</tr>
<tr>
<td>5</td>
<td>20.5</td>
<td>22.5</td>
<td>20.5</td>
<td>19.0</td>
<td>14.0</td>
<td>16.0</td>
<td>15.5</td>
</tr>
<tr>
<td>6</td>
<td>19.0</td>
<td>19.5</td>
<td>16.0</td>
<td>25.0</td>
<td>17.5</td>
<td>19.5</td>
<td>14.0</td>
</tr>
<tr>
<td>7</td>
<td>19.5</td>
<td>20.0</td>
<td>16.0</td>
<td>18.0</td>
<td>19.5</td>
<td>20.5</td>
<td>19.5</td>
</tr>
<tr>
<td>8</td>
<td>13.5</td>
<td>15.5</td>
<td>22.5</td>
<td>18.0</td>
<td>18.5</td>
<td>25.0</td>
<td>24.5</td>
</tr>
<tr>
<td>9</td>
<td>27.0</td>
<td>29.0</td>
<td>25.0</td>
<td>23.5</td>
<td>24.5</td>
<td>30.5</td>
<td>25.5</td>
</tr>
<tr>
<td>10</td>
<td>18.0</td>
<td>28.0</td>
<td>29.0</td>
<td>27.0</td>
<td>12.0</td>
<td>14.0</td>
<td>12.0</td>
</tr>
<tr>
<td>11</td>
<td>41.0</td>
<td>32.0</td>
<td>32.5</td>
<td>35.5</td>
<td>20.5</td>
<td>25.5</td>
<td>37.5</td>
</tr>
<tr>
<td>12</td>
<td>12.0</td>
<td>11.5</td>
<td>11.0</td>
<td>9.5</td>
<td>8.0</td>
<td>8.5</td>
<td>7.0</td>
</tr>
<tr>
<td>13</td>
<td>6.0</td>
<td>5.5</td>
<td>4.0</td>
<td>4.5</td>
<td>4.5</td>
<td>5.5</td>
<td>6.0</td>
</tr>
<tr>
<td>14</td>
<td>10.0</td>
<td>9.5</td>
<td>11.0</td>
<td>10.5</td>
<td>8.5</td>
<td>7.5</td>
<td>9.0</td>
</tr>
<tr>
<td>15</td>
<td>6.5</td>
<td>5.5</td>
<td>6.0</td>
<td>4.5</td>
<td>7.5</td>
<td>4.5</td>
<td>6.0</td>
</tr>
<tr>
<td>16</td>
<td>15.5</td>
<td>10.5</td>
<td>13.5</td>
<td>21.0</td>
<td>14.5</td>
<td>17.0</td>
<td>17.0</td>
</tr>
<tr>
<td>17</td>
<td>31.5</td>
<td>38.0</td>
<td>31.0</td>
<td>21.0</td>
<td>26.0</td>
<td>14.5</td>
<td>12.0</td>
</tr>
<tr>
<td>18</td>
<td>21.0</td>
<td>19.0</td>
<td>15.0</td>
<td>23.5</td>
<td>22.5</td>
<td>14.5</td>
<td>12.0</td>
</tr>
<tr>
<td>19</td>
<td>17.5</td>
<td>7.5</td>
<td>8.5</td>
<td>19.5</td>
<td>18.5</td>
<td>32.0</td>
<td>34.0</td>
</tr>
<tr>
<td>20</td>
<td>18.0</td>
<td>18.5</td>
<td>15.0</td>
<td>20.5</td>
<td>22.0</td>
<td>19.5</td>
<td>17.0</td>
</tr>
<tr>
<td>21</td>
<td>34.5</td>
<td>37.5</td>
<td>40.0</td>
<td>35.5</td>
<td>16.5</td>
<td>11.5</td>
<td>14.0</td>
</tr>
<tr>
<td>22</td>
<td>29.0</td>
<td>31.5</td>
<td>29.5</td>
<td>30.0</td>
<td>24.5</td>
<td>30.5</td>
<td>26.0</td>
</tr>
<tr>
<td>23</td>
<td>25.5</td>
<td>24.0</td>
<td>35.0</td>
<td>30.5</td>
<td>21.5</td>
<td>23.5</td>
<td>28.5</td>
</tr>
<tr>
<td>24</td>
<td>18.5</td>
<td>26.0</td>
<td>17.0</td>
<td>16.5</td>
<td>15.0</td>
<td>11.0</td>
<td>14.0</td>
</tr>
</tbody>
</table>
normal distribution of the values is shown in Fig 4. Normality according to the Kolmogorov-Smirnov test was 0.105 ($P < .05$); skewness was 0.365. The extent values were 4.0 to 41.0 N/cm$^2$ (75% of the values between 10.0 and 30.0 N/cm$^2$).

**Reliability Study**

The data on reliability for fixation detection by the students with respect to the experienced chiropractor were compiled. Note that some student examiners were not compiled: they were lacking certainty in their findings, and the research protocol did not allow them to repeat their palpation. Table 2 represents the compilation of all student examiners. Table 3 exhibits the compilation when examiners exerted an average pressure of palpation, that is to say, 19.5 N/cm$^2$ ± a standard deviation of 8.6 N/cm$^2$. Tables 4 and 5 represent the compilations when the force applied

### Table 2. Reliability results for all examiners combined

<table>
<thead>
<tr>
<th>Number of agreements on presence</th>
<th>Disagreements</th>
<th>Student examiner: absence</th>
<th>Experienced examiner: presence</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Q = 21$</td>
<td></td>
<td></td>
<td>$R = 4$</td>
</tr>
<tr>
<td>Disagreements</td>
<td>$S = 8$</td>
<td></td>
<td>$T = 103$</td>
</tr>
<tr>
<td>Student examiner: presence</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experienced examiner: absence</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Percent agreement = 91.2%. Concordance ($\kappa$) = 0.724.

### Table 3. Reliability results for examiners with average pressure of palpation (10.0-30.0 N/cm$^2$)

<table>
<thead>
<tr>
<th>Number of agreements on presence</th>
<th>Disagreements</th>
<th>Student examiner: absence</th>
<th>Experienced examiner: presence</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Q = 14$</td>
<td></td>
<td></td>
<td>$R = 2$</td>
</tr>
<tr>
<td>Disagreements</td>
<td>$S = 6$</td>
<td></td>
<td>$T = 65$</td>
</tr>
<tr>
<td>Student examiner: presence</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experienced examiner: absence</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Percent agreement = 90.8%. Concordance ($\kappa$) = 0.723.

### Table 4. Reliability results for examiners with high pressure of palpation (30.0-41.0 N/cm$^2$)

<table>
<thead>
<tr>
<th>Number of agreements on presence</th>
<th>Disagreements</th>
<th>Student examiner: absence</th>
<th>Experienced examiner: presence</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Q = 3$</td>
<td></td>
<td></td>
<td>$R = 1$</td>
</tr>
<tr>
<td>Disagreements</td>
<td>$S = 1$</td>
<td></td>
<td>$T = 20$</td>
</tr>
<tr>
<td>Student examiner: presence</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experienced examiner: absence</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Percent agreement = 92.0%. Concordance ($\kappa$) = 0.701.
Table 5. Reliability results for examiners with low pressure of palpation (4.0-10.0 N/cm²)

<table>
<thead>
<tr>
<th>Number of agreements on presence</th>
<th>Disagreements</th>
<th>Student examiner: presence</th>
<th>Experienced examiner: absence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q = 4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disagreements</td>
<td></td>
<td>Student examiner: absence</td>
<td>Experienced examiner: presence</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R = 1</td>
<td></td>
</tr>
<tr>
<td>S = 1</td>
<td></td>
<td>Number of agreements on</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>presence</td>
<td>absence</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T = 18</td>
<td></td>
</tr>
</tbody>
</table>

Percent agreement = 91.7%. Concordance (κ) = 0.748.

was relatively high (above 30.0 N/cm²) and relatively low (below 10 N/cm²). Q values are small, and κ is unstable in Tables 4 and 5.

**DISCUSSION**

Motion palpation is subject to controversy within the chiropractic community. Preliminary reports on the reliability of the test presented interesting percent agreement between examiners. Unfortunately, the percent agreement is not sufficient to declare a test reliable. Subsequent studies using more appropriate statistical analysis by testing the agreement with the coefficient of concordance κ often revealed a weak to fair reliability of the test. The difficulties inherent to the experimental verification of reliability were the subject of several citations in many studies, in particular, the difficulties in standardizing the technical aspects of the test.

More recent studies showed the benefit of conducting the test with a standardization of kinematics. The reliability of the test was largely improved by a standardized method. The results of the current study show once more that the agreement among the examiners is enhanced when the kinematics of the test are reproduced successfully by examiners (κ = 0.724).

Another parameter of importance of the test is its kinetics: the force applied when palpating (pressure of palpation). The current study proposed to verify, by direct measurement, this variable of the test. Our study revealed that for the test in cervical spine rotation, the pressure of palpation ranges between 4.0 and 41.0 N/cm² (averaging near 20 N/cm²). For 75% of the examiners, the pressure of palpation was from 10.0 to 30.0 N/cm², or a pressure applied of 2.0 to 6.0 N for an effective surface of 0.2 cm². By comparison, the average pressure applied to the sensors by the examiners in our study was approximately equivalent to a weight of 400 g or less than 1 lb. Although some pressure was also distributed outside the perimeter of the sensors, this was not measured. For the purpose of our study, the surface of the sensors was considered the useful surface of palpation.

This study shows that selected examiners have a strong interexaminer reliability despite the pressure of palpation ranging from 4.0 and 41.0 N/cm². However, no strong conclusion is given as to the reliability of examiners in lower or higher pressure range, as Q values are small and κ unstable (Tables 4 and 5). We attribute the strong reliability to the fact that the kinematics were highly standardized and to the selection of patients with a history of biomechanical neck pain. We consider the possibility that examiners using pressures of palpation that would deviate significantly from the values of this study could obtain less favorable results than those obtained by our examiners.

In view of this study, the kinematics of the movement appear to be a variable of greater importance than the pressure of palpation, whereas appreciable variations (ranging from 4.0 to 41.0 N/cm²) do not significantly alter the reliability of the test.

The sensors somewhat prevented direct skin contact between the patient’s neck and examiner’s hand during palpation, yet the examiners commented that the presence of sensor and equipment did not obstruct their palpation. These examiners did not seem to rely on palpatory features such as cutaneous temperature or skin texture, but rather relied on proprioceptive clues such as movement or the resistance to it. The clinical judgment, to which motion palpation is subjected, is most probably derived from examiners’ mechanoreceptors and proprioceptors. These still remained functional after the installation of the sensors.

Although we realize that good reliability does not imply validity of clinical decisions, it is recommended, in the light of these results, that the palpatory pressure applied during motion palpation be comfortable for the patient and sufficient for the examiner to determine the presence of fixations while appreciating the dynamics of the spine.

**CONCLUSION**

The pressure applied during palpation as a parameter of the test of motion palpation does not seem to represent a variable relating to interexaminer reliability. Strong reliability of the test was obtained by using subjects with a history of mechanical neck disorders and by standardizing the kinematics of the test. The pressure applied during motion palpation for cervical spine rotation is light to moderate but always comfortable for the patient. Applied pressures can vary tenfold (4-41 N/cm²) and remain reliable provided the kinematics of the test are standardized. Additional research on forces applied during the test of motion palpation will be necessary to better document the results obtained in this study.
ACKNOWLEDGMENTS

We acknowledge the notable contribution of Dr P. Kogon for translating the original manuscript into English.

REFERENCES

20. Faye LJ. Motion palpation of the spine. Huntington Beach (Calif): Motion Palpation Institute; 1981.
CONCURRENT VALIDITY OF FLEXICURVE INSTRUMENT MEASUREMENTS: SAGITTAL SKIN CONTOUR OF THE CERVICAL SPINE COMPARED WITH LATERAL CERVICAL RADIOGRAPHIC MEASUREMENTS

Deed E. Harrison, DC, a Jason W. Haas, DC, b Rene Cailliet, MD, c Donald D. Harrison, PhD, DC, MSE, d Burt Holland, PhD, e and Tadeusz J. Janik, PhD f

ABSTRACT

Objectives: The aim of this study was to compare flexicurve surface contour measurements of the cervical spine with radiographic measurements of cervical lordosis.

Methods: One examiner evaluated 96 patients with chronic neck pain in neutral posture using a flexible ruler, flexicurve, to measure sagittal contour of the skin over the cervical spine from the external occipital protuberance to the vertebra prominens. The flexicurve skin contour and neutral lateral radiographs were digitized and compared. The flexicurve and radiographs were categorized into height-length ratio, curve angle, curve depth, sum of depths, modified Ishihara’s index, and inverse of radius. Mean values, SDs, mean differences, and limits of agreement were calculated. The differences between flexicurve measurement mean values and x-ray mean values were deemed significant if the lower limit of agreement exceeded 15% of the mean values for the x-ray measurements.

Results: For all variables, except the height-length ratio, the mean values of the flexicurve variables differed significantly from the corresponding mean values of the radiographic measurements. All Pearson correlation coefficients were in the very poor range ($r < 0.15$).

Conclusion: The flexicurve sagittal skin contour measurement has poor concurrent validity compared with established radiographic measurements of the cervical lordosis. The flexicurve tracings always predicted lordosis, overestimated the lordosis compared with x-ray values, and cannot discriminate between radiographic lordosis, straightened, S curves, and kyphotic alignments of the cervical curve. (J Manipulative Physiol Ther 2005;28:597-603)

Key Indexing Terms: Radiography; Reproducibility of Results; Flexicurve; Surface Contour

Neck pain affects as many as 70% of individuals throughout their lives.1 A large portion of health care in the United States and throughout the world is devoted to neck problems.2 Besides pain and quality of life outcomes, cervical range of motion3-5 and alignment of the cervical lordosis6-17 have been considered objective outcome measures of patient conditions. Concerning the cervical lordosis, studies have identified a relationship between an abnormal cervical lordosis (straightened and/or kyphotic alignment) and conditions such as headaches,6,7 whiplash-associated disorders,10,11 spondylotic myelopathy,12,13 and degenerative disorders of the disks and vertebral bodies.10,14,15 Furthermore, in the sagittal plane, alteration in the static cervical spinal alignment alters the dynamic kinematics of the cervical spine during routine activities such as flexion and extension.17

Although radiography is considered by some to be a criterion standard for evaluation of both the cervical lordosis6-16 and segmental kinematics during flexion and extension,3,4,17 some investigators are attempting to find...
less-invasive conservative methods for diagnosis of pathological/abnormal spinal conditions. Some researchers have used skin surface contour to analyze a patient’s sagittal plane alignment and have made clinical judgments according to these skin contours.18-22

Investigators have used the flexicurve ruler and similar noninvasive instruments in the measurements of the thoracic and lumbar spines for static position and dynamic motion.18-22 There may be advantages to using surface contour to measure patients’ spinal alignment, such as simplicity, cost-effectiveness, and no radiation exposure. However, past studies comparing surface contour with spinal position have found mixed results concerning reliability and validity of surface assessment compared with radiography.18-24 In at least two studies, surface contour alignment of the lumbar lordosis has not been shown to correlate with measurements of lordosis using lumbar radiographs of the same subject.23,24 Despite this disparity, some authors continue to use the skin contour tools to make measures of sagittal spine alignment in patients and claim that these give valid information.25,26

Only two investigations have been performed comparing the surface contour of the skin overlying the cervical spine with radiographic measurements of the cervical lordosis in the same subjects. Neither of these studies performed a complete comparison of multiple measurement variables between surface contour and x-ray, and neither used the flexicurve.27,28 In light of this, continued investigations into the validity of surface contour measurements should be performed.

This study compared measurements of the flexicurve skin surface contour of the cervical spine to measurements of the cervical lordosis on digitized lateral cervical radiographs. Because previous investigations had reported moderate validity of the surface contour of the thoracic spine compared with x-ray measurements of thoracic kyphosis,20 we decided to investigate the concurrent validity of flexicurve measurements of the surface contour of the neck region with measurements of cervical lordosis on lateral radiographs. We hypothesized that the flexicurve skin contour would be a valid representation of the cervical lordosis compared with x-ray measurements.

METHODS

Ninety-six consecutive patients with chronic neck pain, who presented for treatment of this condition at a spine clinic in Elko, Nev, were evaluated by one examiner using two instruments: the flexicurve ruler and traditional lateral cervical radiography. Subjects reviewed the institutional review board–approved study protocol and provided informed consent for their participation. The subjects consisted of 60 females and 36 males whose mean age was 40.1 ± 17.9 years. The average weight was 78.9 ± 20.7 kg, and the average height was 168.7 ± 12.7 cm. Their average Numerical Rating Scale scores were a mean of 4.6 ± 2.5 (0 being no pain and 10 being worst pain ever/bedridden).

The patients were instructed to stand in front of the x-ray cabinet, in a comfortable neutral position, and place their right shoulder against the grid cabinet. The examiner then instructed the patients to close their eyes, nod their head up and down two times, come to a position in which they felt their head was neutral, and then open the eyes without moving their position. This neutral posture position of the head and neck posture is repeatable.29,30 They were told that they would feel some pressure on the posterior portion of their neck as the examiner placed the flexicurve and that they should not move their head from the neutral position.

The flexicurve instrument is a drafting device, which is a flexible metal strip covered with plastic marked with a metric ruler, and can be purchased at any variety store. This device deforms in only one plane and retains the shape of deformation. One examiner (JWH) placed the bottom end of the flexicurve instrument at the vertebra prominens (VP) and then molded the instrument to fit the posterior aspect of the patients neck until reaching the inferior aspect of the external occipital protuberance (EOP) (Fig 1). The VP and EOP landmarks were chosen as the beginning and terminal ends of the curve because they are easily identifiable and they are a general demarcation between the cervical and thoracic spine.

**Fig 1.** The flexicurve was positioned along the posterior skin contour in the cervical region from the EOP to the VP. The contour of the flexicurve was traced onto paper and then digitized.
The examiner (JWH) had used the flexicurve for several months and was proficient in its use. After approximating the instrument to the patient’s neck, the examiner created a paper template with the shape of the flexicurve. Immediately after flexicurve measurement, the patient was exposed to a lateral cervical radiograph taken at 182.9 cm (72 in) by the same examiner. The patient retained the same neutral position of the cervical spine during lateral cervical x-ray.

Before any data analysis, the same examiner (JWH) divided the flexicurve tracings into 6 arcs by first choosing the apex of the curve as the largest perpendicular distance from a straight line through the top (EOP) and bottom (VP) of the tracing. Next, this straight line was used to divide the flexicurve tracing into thirds between the bottom (VP) and the point used to determine the apex, and additional thirds between the point used to determine the apex and the top (EOP) (Fig 2).

A separate examiner (DEH) digitized 7 points on each flexicurve tracing. These points included (1) top of the tracing (EOP), (2) one third from the top to the apex, (3) two thirds from the top to the apex, (4) apex, (5) one third from the apex to the bottom, (6) two thirds from the apex to the bottom, and (7) the bottom of the curve (VP). All 96 flexicurves were digitized in random order. The same examiner (DEH) then digitized all 96 lateral cervical radiographs in random order. After the data were digitized, the data were analyzed by a different examiner (TJJ) using a system composed of a GP-9 sonic digitizer (Science Accessories Corporation, Shelton, Conn) and a personal computer. As reported by the manufacturer, the resolution of the sonic digitizer is 0.125 mm, but the average accuracy of digitized x-ray points is 0.5 mm in clinical use. The computational results were statistically analyzed by a different examiner (BH).

On the flexicurve tracing, the straight line distance (VP to EOP) is designated the height, whereas the length is the curvilinear distance along the curve (7 digitized points). Both the flexicurve tracings and the lateral cervical radiographs were measured for height-length (H/L) ratios, mean values, SDs, radius of curvature, angle of curvature, depth at the top one-third and two-third points of the curve, depth at the apex, and depth at the lower one-third and two-thirds points of the curve. When summing depths of the cervical lordosis and dividing by the height of the column on the radiograph, it is termed the modified Ishihara’s index. From the digitized points along the flexicurve tracing, a best-fit circle was created, and using the center of this circle, the angle of curvature and radius of curvature were calculated.

A third examiner (DDH), not privy to the outcome of the digitization process, classified the radiographs by the shape of the sagittal cervical spine into lordotic, military, S shape, and kyphotic. This classification of curves has been previously discussed and reported to be reliable. Segmental and global angles of the cervical curve were determined with the posterior tangent method, which has been shown to be reliable. For the posterior tangent method, lines are drawn on the posterior body margins of C2 through C7. Segmental angles (relative rotation angle) are created at each juxtapositioned pair of posterior tangents, and a global angle of curvature (absolute rotation angle) is measured between the posterior tangents on C2 and C7.

The limits of agreement approach is recommended by Bland and Altman for comparing a new measurement method with an existing measurement method. We used this approach to compare the flexicurve and x-ray measurements. All calculations were done using S-Plus 6.1 (Insightful Corp, Seattle, Wash). For reasons given by Bland and Altman, the use of Pearson correlation coefficients is inappropriate for this problem. However, because 2 recent articles on the surface contour of the cervical spine used Pearson correlation coefficients, we provided these for comparison.

All variables, except the radius (R) and the H/L ratio, were found to have normal distributions as required by the Bland and Altman approach. Because the reciprocal of the radius, 1/R, was found to be approximately normally distributed, we used this variable instead of R itself. (A finding that two
Table 1. Statistical results of 5 measurements made on lateral cervical radiographs and flexicurve posterior neck skin contour on 96 subjects

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean_X</th>
<th>SD_X</th>
<th>Mean_F</th>
<th>SD_F</th>
<th>DiffMean</th>
<th>15% Mean_X</th>
<th>LLA</th>
<th>ULA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arc angle</td>
<td>20.0</td>
<td>15.1</td>
<td>82.3</td>
<td>20.9</td>
<td>62.9</td>
<td>3.1</td>
<td>15.1</td>
<td>110.8</td>
</tr>
<tr>
<td>AP depth</td>
<td>5.5</td>
<td>4.4</td>
<td>22.2</td>
<td>5.7</td>
<td>16.7</td>
<td>0.8</td>
<td>3.1</td>
<td>30.4</td>
</tr>
<tr>
<td>Depth sum</td>
<td>21.8</td>
<td>18.0</td>
<td>82.5</td>
<td>20.7</td>
<td>60.7</td>
<td>3.3</td>
<td>9.6</td>
<td>111.8</td>
</tr>
<tr>
<td>Ishihara</td>
<td>.20</td>
<td>.16</td>
<td>.72</td>
<td>.18</td>
<td>.53</td>
<td>.03</td>
<td>.07</td>
<td>.99</td>
</tr>
<tr>
<td>Radius-inv</td>
<td>.0031</td>
<td>.0024</td>
<td>.012</td>
<td>.003</td>
<td>.0087</td>
<td>.0005</td>
<td>.0014</td>
<td>.016</td>
</tr>
<tr>
<td>HtoLF</td>
<td>.0031</td>
<td>.0024</td>
<td>.012</td>
<td>.003</td>
<td>.0087</td>
<td>.0005</td>
<td>.0014</td>
<td>.016</td>
</tr>
<tr>
<td>AngleF</td>
<td>.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ApexDF</td>
<td>.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SumDX</td>
<td>.13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IshiF</td>
<td>.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RinvF</td>
<td>.11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| 1/R measurements are comparable is equivalent to the finding that two R measurements are comparable.)

Because radiographic measurements are the gold standard, a clinical decision was made to use 15% of x-ray mean values (Mean_Xs) for determining if the flexicurve mean values (Mean_Fs) are acceptably close to x-ray values.

RESULTS

Variables comparing the flexicurve and radiographs were categorized into H/L ratio, angle of curvature, curve depth, sum of the depths, modified Ishihara’s index, and the inverse of the radius (1/R). Table 1 displays the mean values, SDs, mean differences, and limits of agreement for each of these variables.

In accordance with Bland and Altman, the differences between Mean_Fs and Mean_Xs were deemed significant if the lower limit of agreement exceeded 15% of Mean_X. By this standard, for all variables (except H/L, which was not normal distributed), Mean_F differed significantly from Mean_X.

Because H/L ratio was not normally distributed, the Bland and Altman approach cannot be used for H/L. For a different analysis, it was noted that as the H/L ratio approached 1.0, the x-ray measurements and flexicurve measurements began to agree. It was also noted that the flexicurve skin contour always appeared lordotic (with an anterior to posterior depth). Thus, he shape of the sagittal cervical curvature (lordotic, military, S shape, kyphosis) was compared with the flexicurve shape for H/L ratios close to 1.0, that is, H/L_F is greater than 0.96 and less than 1.0. Of 7 subjects in this range, although the flexicurve predicted lordosis for all subjects, only two subjects had a cervical lordotic shape on x-ray. In fact, only 55 of 96 subjects had a lordotic shape on the radiographs, whereas the flexicurve depicted all subjects with a lordosis.

For comparison to previous studies, we calculated Pearson correlation coefficients for all variables except radius of curvature (Table 2). When comparing like variables for the flexicurve and radiographic measurements, all Pearson r’s were between 0.10 and 0.15. This is in the very poor range.

Segmental angles of curvature were measured on the radiographs. Table 3 provides the mean values and SDs for segmental angles of curvature from C2 to C7 obtained from the posterior tangent method. Segmental angles cannot be obtained from the flexicurve.

DISCUSSION

The flexicurve skin contour and lateral cervical x-ray measurements were compared for 96 neck pain subjects. Although the flexicurve predicted lordosis for all 96 subjects, only 55 had a lordotic configuration on the lateral cervical radiograph. From Table 1, the flexicurve consistently predicted an increase in cervical curvature measures compared with x-ray measures. We speculate that this is due to at least two factors: the chosen end points of our flexicurve measures (EOP and VP) and the reduction in spinous process length in the midcervical region (C3-C6). With regard to our use of the EOP, we used the superior aspect of this bony point and can identify no comparably reproducible landmark to use.

The variables (angle of curvature, curve depths, sum of depths, radius of curvature, and Ishihara’s index) had ranges of the lower limit of agreement in excess of 15% of Mean_X
(the mean of any x-ray measurement), thus indicating poor agreement between the flexicurve and radiographic measurements. In addition, all Pearson r values were in the very poor range (r < 0.15).35,36

Some limitations of this study might be slight positional changes on the flexicurve paper drawings when pressure is applied to the patient’s neck while placing the flexicurve over the skin or in tracing the flexicurve contour on a sheet of paper. A future method of scanning skin contour could minimize these effects. Second, it might be thought that postural positioning, x-ray positioning, and radiographic analysis provide confounding variables that influence results. However, all these factors have been shown to be highly repeatable and have negligible effects on angles and distances measured.30 Also, obesity, muscular development, previous trauma, and generally complex biomechanics of the cervical spine may account for some differences in the measurement of surface contour compared with lateral radiographs.

Another concern might be our choice of the Harrison posterior tangent and Ishihara’s index methods for quantification of cervical lordosis on x-ray compared with similar measurements of the flexicurve skin contour tracings. In the recent literature, there are 4 primary radiographic line drawing methods for measurement of cervical lordosis that have reliability, normative data, and predictive validity (ability to discriminate between healthy and abnormal patient populations). These 4 methods are (1) Harrison posterior tangent method,8,11,16,31 (2) Ishihara’s index,7,32 (3) the area under the curve,12 and (4) Cobb method.9,14,31

In the current study, we choose the Harrison posterior tangent and Ishihara’s index because these fit two criteria: the method must be established and readily available for clinical use, and second, the method must be able to be performed in a similar manner on the flexicurve tracing. The area under the curve12 could have been calculated; however, we performed radius of curvature and angle of curvature, which is similar and more clinically relevant. Cobb method9,14,31 was not chosen because this is a cross section through the upright curved column using the endplates and there is no way of reproducing this on the flexicurve tracing.

Furthermore, the Harrison posterior tangent method and the Ishihara’s method (CCI) are able to be converted to each other with the equation: CCI = −0.06 + 0.71 × C2-C7 posterior body angle. The reason that these can be converted is that they both depend upon the curvature of the posterior bodies of the cervical lordosis. In fact, if the deflection of a column is represented by a polynomial formula derived from digitizing posterior vertebral body margins, then it is known in engineering that the first derivative is the slope (posterior tangent lines), the second derivative is the bending moment, and the third derivative is the shear. In vertically loaded columns such as the spine, the second derivative is related to the bending moment through the radius of curvature.37 Thus, the posterior tangents, as slopes or first derivatives, are a basic part of an engineering analysis of deflected columns. Thus, the posterior bodies are the best way to analyze the curve of the cervical spine and to compare the flexicurve (tangent to the skin) with the radiographic angle of cervical lordosis.

In addition, although clinicians are familiar with correlation coefficients, the limits of agreement of Bland and Altman34 are less easily understood. Bland and Altman34 stated, “How far apart measurements can be without causing difficulties will be a question of judgment.” Therefore, some might question our clinical criterion of using 15% of Mean_Xs as a guideline. However, the facts are (1) using 35% of Mean_Xs instead of 15% will give the same poor validity results; (2) the flexicurve cannot predict radiographic configurations of the sagittal cervical spine found on the x-ray; (3) the flexicurve cannot provide segmental and global angles of lordosis between C2 and C7; and (4) the wide range between the upper and lower limits of agreement for all 5 variables in Table 1 indicates large differences between flexicurve variables and radiographic variables.

We were able to locate two other studies in literature that compared cervical skin contour with the lateral cervical radiograph.27,28 Using Pearson product-moment correlation, Refshauge et al28 compared the surface contour of 24 healthy subjects to lateral cervical x-rays. A 40-mm metal strip was placed on the skin overlying the spinous processes of C2, C4, and C7. On the radiograph, the vertebral body centroid of C2, C4, and C7 was located and used to create two angles of curvature: cervical inclination (C2-C7 centroids compared with horizontal) and cervical angle (C2-C4 centroids compared with C4-C7 centroids). Using the same two angles of surface contour, Refshauge et al28 found poor to moderate correlation (r = 0.55 and 0.65) for the two angles of x-ray and surface contour. In 34 healthy young women, Johnson27 used simple correlation coefficients to determine associations between 4 cervical angles of x-ray alignment of the upper cervical spine (C1-C4) and the angle between C7 spinous

<table>
<thead>
<tr>
<th>Table 3. Segmental angles (RRAs) of cervical curvature obtained from applying the posterior tangent method to the radiographs of 96 pain subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Segmental angles from intersecting tangents on the posterior body</td>
</tr>
<tr>
<td>RRA C2-C3</td>
</tr>
<tr>
<td>Measurement</td>
</tr>
</tbody>
</table>

Values are in degrees and are presented as mean ± SD. The flexicurve cannot provide such an analysis.
and the tragus of the ear compared with horizontal (craniocervical angle). No significant correlations between the craniocervical angle and the x-ray alignment of C1 through C4 were found.

Although the current article results are consistent with the findings of Refshauge et al.\cite{28} and Johnson,\cite{27} we believe our methods and analysis to be an improvement and provide more details. First, as stated, according to Bland and Altman,\cite{34} Pearson product-moment correlation coefficients are generally considered poor indications of agreement for two methods of measurement. Accordingly, we used the ranges between the lower and upper limits of agreement for each variable as suggest by Bland and Altman.\cite{34} Second, we used a large number of subjects (96) and included only patients with neck pain as would be the case in clinical practice. Lastly, we included a more detailed analysis of 6 measurements of the entire sagittal cervical spine.

Validity refers to the appropriateness, legitimacy, truthfulness, or effectiveness of a particular topic. Crocker\cite{38} has defined validity as “the extent to which measurements are useful for making decisions relevant to a given purpose.” There are multiple types of validity, including face, construct, content, concurrent, and predictive validity. Using these types of validity, our analysis of the flexicurve surface assessment for cervical curvature indicates that concurrent validity is poor. The flexicurve and the radiographic measurements cannot be interchanged. Thus, the flexicurve is not a valid representation of cervical curvature because this measure cannot be used to decide the angle of cervical lordosis and the shape of cervical lordosis, and decide whether a given patients cervical curvature falls outside the normal range.\cite{7,8,16}

As a final comment on the flexicurve instrument, there are published normal values for the cervical lordosis as measured via radiography.\cite{9} Unfortunately, nowhere was it found in literature normal values for the surface contour of the skin overlying the spine. Thus, without a standard definition of normal values or ranges to compare with, the clinician cannot tell what an abnormal value for the flexicurve measurement might be.

**CONCLUSION**

The results indicate that the flexicurve sagittal skin contour measurement has poor concurrent validity when compared with radiographic measurements of the cervical lordosis. The flexicurve tracings always predicted lordosis, overestimated the lordosis compared with x-ray values, and cannot discriminate between radiographic lordosis, straightened, S curves, and kyphotic alignments of the cervical curve. Because of the disparity between cervical surface contour and neutral lateral radiographs, it is recommended that clinicians should not base outcomes for cervical spine lordosis on the flexicurve surface contour analysis. The flexicurve does not show the validity to replace the radiography as a “gold standard” for evaluation of cervical lordosis. Future work remains to establish whether the flexicurve instrument can provide any meaningful outcome of sagittal cervical contour.

**REFERENCES**

16. Harrison DD, Janik TJ, Troyanovich SJ, Holland B. Comparisons of lordotic cervical spine curvatures to a
VALIDATION OF THE SPIN-T GONIOMETER, A CERVICAL RANGE OF MOTION DEVICE
Shabnam Agarwal, MSc,a Garry T. Allison, PhD, PT,b and Kevin P. Singer, PhD, PTc

ABSTRACT

Objective: To test the validity of the Spin-T goniometer for the assessment of cervical range of movement.
Methods: A linear regression analysis for paired neck movements using first a foam head model and then human subjects was performed to quantify the differences between the measurements obtained from the MotionStar, a movement-tracking device, and the Spin-T. A within-subject repeated measures design using simultaneous data acquisition was completed.
Results: The coefficient of determination ($R^2$) for all planes of cervical range of motion for both model and human data sets was higher than 0.99. The regression equations for the model data showed no significant ($P > .05$) intercept for flexion-extension and lateral rotation. Human data showed statistically significant intercept for flexion-extension (mean, $-0.52^\circ$) and lateral flexion (mean, $0.81^\circ$) at $P < .05$.
Conclusion: This study quantifies the difference between the MotionStar and the Spin-T goniometer and documents the systematic error between the measures. Where the error reached statistical significance, the magnitude of the error was very small ($<1.5^\circ$). The results of this study suggest that the Spin-T goniometer may be used as a valid measuring instrument for cervical range of movement. (J Manipulative Physiol Ther 2005;28:604-609)

Key Indexing Terms: Reproducibility of Results; Cervical Vertebrae; Range of Motion; Articular; Validity

The classic spinal motions are flexion, extension, lateral rotation, and lateral flexion. Cervical spine movement is difficult to investigate accurately because of its anatomic structure and individual compensatory movements that may be associated with habit, posture, or pain. Motion in the cervical spine may be divided into the upper cervical spine (occiput to C2) and the lower cervical spine (C3 through T1). Movements of the upper cervical spine include flexion-extension and lateral rotation with minimal lateral flexion, whereas in the lower cervical spine, all 4 movements occur.1 Movements in the cervical spine are determined by the orientation of the facets, passive tension of the ligaments, muscles, joint capsule, and fibers of the anulus fibrosus.1

Normal variation of the cervical range of motion (CROM) is influenced by age and sex,2-4 degeneration, pathology, surgery, or trauma, as well as factors such as pain,5,6 muscle spasm, and whether the movement is performed actively or passively.

A subjective, qualitative observation of the range and path of motion is normally performed by clinicians to analyze passive and active movements. Lack of convenient, valid, and reliable instruments may be a reason why measuring instruments are not used in routine clinical practice. Measuring instruments may be time-consuming for the operator and cumbersome for the patient. Decisions regarding intervention and treatment are often based partly on joint motion, and clinicians need to justify their choice of treatment modality based on an objective assessment of the CROM. Many different methods and instruments have been used to assess CROM. Validity of measuring equipment has been reported less frequently than reliability.

Concurrent validity is established by comparing test scores with a recognized gold standard, the criterion. If a high concurrent validity is established, then clinical utility is related to the measurement sensitivity and the ease and logistics of the clinical tool in the normal physiotherapy, and rehabilitation setting is considered.

a Chief Physiotherapist, Belle Vue Clinic, Kolkata, India.
b The Centre of Musculoskeletal Studies, Perth, Western Australia.
c The Centre of Musculoskeletal Studies, Perth, Western Australia.

Sources of support: No external funds were received for this research.
Submit requests for reprints to: Shabnam Agarwal, MSc, Belle Vue Clinic, 9, Dr. U.N. Brahachari St., Kolkata 700 017, India (e-mail: shabnamagarwal@vsnl.net).
0161-4754/$30.00
Copyright © 2005 by National University of Health Sciences.
doi:10.1016/j.jmpt.2005.08.015
The CROM device has been used to report concurrent validity of a single inclinometer. The inclinometer was validated for flexion-extension (ICC = 0.80) and lateral flexion (ICC = 0.79), but not for rotation (ICC = −0.18). A study by Haynes and Edmondston showed that the CROM device could not accurately measure natural composite rotation movements. The aim was to establish if the Spin-T and the CROM device could accurately measure natural rotation movement. The devices were placed on a testing instrument which could be positioned at preset angles of rotation with/without a tilt to mimic the lateral flexion that occurs ipsilaterally to and concomitantly with cervical rotation. The Spin-T correlated positively with the testing instrument for rotation with accompanied tilt up to 15° (ICC > 0.99), whereas the CROM device showed a poor concordance (ICC = 0.50) at rotation with 10° tilt. The concurrent validity of the CROM device has been evaluated against radiographs in the sagittal plane. In flexion, the coefficient of determination was $R^2 = 0.94$, $r = 0.97$ at $P < .001$. The slope value was 0.98 with a y-axis intercept of −0.08. In extension, $R^2 = 0.97$, $r = 0.98$ at $P < .001$ with a slope value of 1 and intercept of −2.1. Radiographs in the flexion-extension view have also been used as a gold standard for a pendulum goniometer. The pendulum goniometer showed a positive correlation ($r = 0.97$) with the radiographs for the entire head on neck motion.

Ultrasound-based motion analyzers have been validated against a precision goniometer and a digital inclinometer. A maximum measurement difference of 0.6° was calculated between the CMS 3D real-time motion analyzer (Zebris Medizintechnik GmbH, Isny, Germany) and the precision goniometer. In clinical terms, a 1° error is acceptable. The CMS 70P ultrasound system (Zebris Medizintechnik GmbH) was found to be accurate in comparison with the digital inclinometer.

Christensen validated the CA 6000 Spine Motion Analyzer (Orthopedic Systems Inc, Union City, Calif) electrogoniometer with two manual protractors. Neck movements in all 6 directions were tested with 4 to 5 recordings measured in each tested direction. The electrogoniometer was not found accurate with the mean difference ranging from 2% to 11.5%. The CA 6000 Spine Motion Analyser is expensive and ideally suited for research laboratories.

Studies that establish concurrent validity between clinical tests of CROM and gold standard criteria determine the degree of concordance between the two measurements. It is the clinician who then uses this information to consider if the magnitude of the variance between the two systems is small enough to justify the use of the clinical tool.

The Spin-T goniometer has been devised to measure composite cervical spine movements. The purpose of this study was to compare measurements of the simple, clinical cervical spine Spin-T goniometer with that of a high-resolution motion tracking system (MotionStar; Ascension Technology Corporation, Burlington, Vt) for CROM in different planes.

METHODS

Subjects

Four male subjects (age range, 28-45 years) with no history of head or neck pain volunteered to take part in this study. The subjects performed cervical spine flexion-extension, lateral bending, and rotation movements. The Spin-T goniometer was used to measure cervical spine movements in flexion, extension, and lateral bending. The MotionStar system was used to track the movements of the sensors placed on the subjects' heads. The subjects performed movements at a comfortable pace, and the movements were recorded for analysis.

The Spin-T goniometer was strapped on the subject’s head. The T square is positioned along the spindle of the flexion-extension dial to provide a perpendicular reference to the wall.

Fig 1. The Spin-T goniometer strapped on the subject’s head. The T square is positioned along the spindle of the flexion-extension dial to provide a perpendicular reference to the wall.

Fig 2. The Spin-T on a foam head model placed in front of a wall. One sensor of the MotionStar can be seen on top of the foam head, whereas the other is placed parallel to it, in front of the foam head. The MotionStar is placed to enable it to track movements of its sensors.
The concurrent validity of the Spin-T goniometer was tested for movement in 3 cardinal planes using the MotionStar 3D position sensor as the gold standard. The comparisons were undertaken using a foam head model and also using control volunteers.

Equipment

**MotionStar.** CROM was assessed using a DC magnetic motion capture system (Ascension Technology Corporation, Burlington, Vt), integrated with purpose-designed software (Labview V5.0, National Instruments, Austin, Tex). The MotionStar tracks the location and movement of one or more sensors in a designated field at approximately 86 Hz, and the Euler angles are transposed to angles in the cardinal planes. The Euler transformation was checked using a triaxial protractor of known accuracy (0.5°).

**Spin-T goniometer.** The Spin-T goniometer\(^\text{12}\) consists of a spectacle-type aluminum frame, positioned on the nose with velcro straps. Three 360° dials (marked at 1° intervals) attached to the frame lie in orthogonal planes reflecting the principal movement planes. An L-shaped rectangular plastic spindle pivots around the center of each dial (Fig 1) with the horizontal portion of the L touching the dial (a red line at one end of the spindle coinciding with the degree markings of the dial along its circumference). The orientation of each dial is referenced and zeroed to the perpendicular plane of the laboratory wall. This is achieved by the use of a lightweight, rigid aluminium T square oriented specifically for each dial (Fig 1). Once the reference position is established, the degrees of relative movements in each plane is assessed by using the T square to reset the spindle on each dial. From this, excursion in that plane is documented.

**Testing Protocol**

The validity of the Spin-T goniometer was tested in 3 cardinal planes using the MotionStar, which tracked and reported composite cervical movement in 3 planes simultaneously. Experimentally, it has been proved that the Spin-T is capable of measuring lateral rotation with concurrent

### Table 1. Comparison of foam model and human regression analysis for movements of the cervical spine

<table>
<thead>
<tr>
<th>Movement</th>
<th>(R)</th>
<th>(R^2)</th>
<th>(df)</th>
<th>Intercept</th>
<th>(t)</th>
<th>(P)</th>
<th>(95%\ CI\ for\ intercept)</th>
<th>(95%\ CI\ for\ slope)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flex-ext (f)</td>
<td>.998</td>
<td>.997</td>
<td>22</td>
<td>−0.43</td>
<td>−1.01</td>
<td>.32</td>
<td>−1.33, 0.45</td>
<td>0.99, 0.66</td>
</tr>
<tr>
<td>Flex-ext (h)</td>
<td>.999</td>
<td>.998</td>
<td>82</td>
<td>−0.52</td>
<td>−3.93</td>
<td>**</td>
<td>−0.79, −0.26</td>
<td>0.98**, 4.50</td>
</tr>
<tr>
<td>Lat flex (f)</td>
<td>1.000</td>
<td>.999</td>
<td>22</td>
<td>−0.35</td>
<td>−2.48</td>
<td>**</td>
<td>−0.64, −0.05</td>
<td>1.02**, −4.60</td>
</tr>
<tr>
<td>Lat flex (h)</td>
<td>.998</td>
<td>.997</td>
<td>76</td>
<td>0.81</td>
<td>4.67</td>
<td>**</td>
<td>0.46, 1.15</td>
<td>1.01, −1.85</td>
</tr>
<tr>
<td>Lat rot (f)</td>
<td>1.000</td>
<td>1.000</td>
<td>22</td>
<td>−0.09</td>
<td>−0.66</td>
<td>.51</td>
<td>−0.39, 0.20</td>
<td>1.01**, −3.00</td>
</tr>
<tr>
<td>Lat rot (h)</td>
<td>1.000</td>
<td>.999</td>
<td>70</td>
<td>−0.11</td>
<td>−1.06</td>
<td>.28</td>
<td>−0.31, 0.09</td>
<td>0.98**, 4.33</td>
</tr>
</tbody>
</table>

Double asterisks indicate significance. Flex, flexion; ext, extension; lat, lateral; rot, rotation; f, foam model; h, human head.

---

Fig 3. The regression plots and equations for dual cervical spine range of motion determination using the MotionStar and Spin-T Instruments on a foam model head. A, Cervical extension-flexion. B, Lateral flexion. C, Lateral rotation. All measurements are in degrees.
lateral flexion reliably and accurately. However, with the design of the Spin-T, it was possible to measure movement in only one plane at one time. Hence, concurrent movements in other planes could not be measured simultaneously. The comparisons were undertaken using a foam head model and also using control volunteers.

Two sensors of the MotionStar were placed parallel to each other, one on the reference table and one on top of the foam head model (Fig 2). The purpose of this was to allow pure movements in specific planes. The model was moved in increments of angles in all 3 reference planes and held in a stable position while readings of the Spin-T goniometer and the MotionStar were recorded simultaneously. A total of 72 paired data sets were recorded during flexion-extension and lateral flexion and rotation movements.

The second step of validity testing was undertaken to reflect the clinical setting. Four men performed a series of incremental CROM tests that covered their available range. A total of 234 readings in flexion, extension, lateral flexion, and lateral rotation for each subject were taken.

### Statistical Analysis

A repeated measures design using a linear regression model was used. Simultaneous data acquisition of CROM was performed for a series of ranges of motion in all cervical movements, namely, flexion-extension and lateral flexion (left and right) and rotation (left and right). Paired data sets of movements in all 6 directions were compared for the MotionStar data and the Spin-T data using, first, a linear regression model analysis and, secondly, 95% limits of agreement. Coefficients of determination ($R^2$) were calculated and confidence intervals (CIs) assessed for systematic change in the intercept. $P < .05$ was adopted as the criterion for accepting statistical differences.

### Results

#### Model Head Comparison

The model head data showed a positive correlation with the MotionStar with $R$ values higher than 0.998 for displacements in all directions. Table 1 shows the regression data for induced movements of the cervical spine using a model head. The regression data illustrates a high coefficient of determination ($R^2 > 0.991$) for individual movements with a mean root mean square error of $1.0^\circ$. Scatter diagrams (Fig 3) for linear regression analysis were constructed for Spin-T vs MotionStar measurements.

$R^2$ values were close to 1. Most of the variation of the measurements obtained with Spin-T compared with the independent variable, the MotionStar, are shown in Fig 3. For all movements, the intercept was not significantly different from zero ($P < .05$) except for lateral flexion, which showed a mean error of $-0.35^\circ$ (95% CI, $-0.64$ to $-0.05$; $P < .05$). The slope values for lateral flexion (1.023) and lateral rotation (1.01) were significantly different from 1. However, the lower and upper 95% CI for both movements included 1, and the difference was small (Table 2).

#### Human Data Comparisons

Spin-T measurements of humans showed a high coefficient of determination ($R^2 > 0.982$) for all discrete neck movements and $R^2$ higher than 0.997 for paired movements.

---

**Table 2. Range of 95% CI for intercept and slope values for model and human measurements**

<table>
<thead>
<tr>
<th></th>
<th>Foam model</th>
<th>Human</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flexion-extension</td>
<td>1.78</td>
<td>0.53</td>
</tr>
<tr>
<td>Lateral flexion</td>
<td>0.58</td>
<td>0.69</td>
</tr>
<tr>
<td>Lateral rotation</td>
<td>0.60</td>
<td>0.41</td>
</tr>
<tr>
<td>Slope</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flexion-extension</td>
<td>0.04</td>
<td>0.01</td>
</tr>
<tr>
<td>Lateral flexion</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>Lateral rotation</td>
<td>0.01</td>
<td>0.01</td>
</tr>
</tbody>
</table>

---

**Fig 4. The regression plots and equations for dual cervical spine range of motion determination using the MotionStar and Spin-T Instruments with 4 subjects. A, Cervical extension-flexion. B, Lateral rotation. C, Lateral flexion. All measurements are in degrees.**
DISCUSSION

Compared with human data, the foam head model measurements were more accurate. Reasons for this difference may be attributed to the fact that while testing humans, there would be inadvertent movement, whereas the head model could be fixed into a stationary position. Another explanation is that the number of head model measurements (n = 72) were fewer compared with human measurements (n = 234). As a result, the 95% CI were reduced and smaller differences detected (Table 2). This is apparent when comparing the similar value intercepts and their respective range of the 95% CI values for flexion-extension (head model: intercept = -0.43°, range = 1.78°; human: intercept = -0.52°, range = 0.53°). Therefore, these results suggest that although the bias is similar, it is the sample size that contributes to the detection of these systematic differences in the human test conditions. However, as human measurements are only slightly less exact compared with human data, the foam head model measurements can be considered accurate when measuring cervical spine movements in human subjects.

Statistical significance may not always translate to clinical relevance. The spread of error, as indicated by the root mean square value, is <2°. The source of this error may be explained by the error in the intercept and slope values. For human measurements, the intercept values for flexion-extension, lateral flexion, and lateral rotation are all within 1.5°. Hence, there is a constant error of only 1.5° between the Spin-T measurements and the MotionStar (underestimation for flexion-extension and lateral rotation and overestimation for lateral flexion), which, although statistically significant, is well within minimal clinical difference considering the natural within-subject variation of CROM.18 This minimal difference of 1.5° may actually arise from the clinician recording to the nearest degree. The MotionStar does not have this particular source of error, although it may be susceptible to error because of distortions in the magnetic field or software interpolation. Hence, it may be concluded that Spin-T measurements reflect the MotionStar to within 1.5° error. However, what is important to note is that these error values are sufficient to detect meaningful clinical changes. The slope explains the linear scaling error according to the actual range. Accounting for the 0.5° constant error for flexion-extension and no constant error for lateral rotation, these sources of error are relatively small compared with routine clinical assessments for CROM.

The Spin-T goniometer fulfills all aspects of criterion and external validity. Criterion validity justifies the validity of the measuring instrument by comparing measurements made to a well established “gold standard” of measurement. Readings from the Spin-T goniometer were simultaneously compared with readings from the MotionStar. The regression statistics showed that there is a high concordance across the full range of movement during different cardinal planes of assessment. The Spin-T also has external validity because it was used on a group of normal subjects without any laboratory conditions, except for the normal standardization procedures. Hence, a similar methodology can easily be replicated in a clinical trial or in routine clinical practice for assessment of CROM.

The Spin-T is portable and easy to arrange. The maximum time required is 3 minutes from sitting the patient in position to removing the instrument from the patient’s head. The Spin-T uses a nearby wall to reference the angle and permits an accurate method of measuring composite cervical movements. The Spin-T has also been proved as a reliable instrument on 23 subjects with a high intrarater reliability (>0.87 and >0.91 for each examiner, respectively) and an interexaminer reliability higher than 0.75 for different neck movements.12

The Spin-T can be used in a clinical trial and, as well, can provide the clinician an efficient method to measure objectively natural cervical movements in a clinical setting. The findings of this study suggest that the Spin-T is a valid form of CROM assessment and, therefore, may provide the clinician with an alternative to more technical and expensive research alternatives such as the electrogoniometer17 and ultrasonography.15

CONCLUSION

The Spin-T goniometer is accurate to within 2° in all planes and ranges when compared with 3D electromagnetic assessments. Hence, the Spin-T goniometer may provide a valid assessment of composite and natural CROM.

REFERENCES

6. Jordan A, Mehlisen J, Ostergaard K. A comparison of physical characteristics between patients seeking treatment for neck pain...


VALIDITY OF THE LATERAL GLIDING TEST AS TOOL FOR THE DIAGNOSIS OF INTERVERTEBRAL JOINT DYSFUNCTION IN THE LOWER CERVICAL SPINE

Cesar Fernández-de-las-Peñas, PT,a Cristobal Downey, PT, MSc,b and Juan Carlos Miangolarra-Page, MD, PhDc

ABSTRACT

Objective: To determine if the lateral gliding test for the cervical spine is a valid clinical test compared with radiological assessment as a tool for the diagnosis of intervertebral joint dysfunctions in the lower cervical spine in patients presenting with mechanical neck pain.

Methods: Twenty-five patients with mechanical neck pain presenting with an asymmetry of at least $5^\circ$ between left and right cervical lateral flexion and diagnosed with an intervertebral joint dysfunction in the lower cervical spine based on the lateral gliding test were studied. Two anterior-posterior x-rays were performed on each patient at maximum end-range of right and left cervical lateral flexion. The intervertebral motion was compared between the hypomobile side and the contralateral side at the level diagnosed as hypomobile by the lateral gliding test.

Results: The asymmetry between left and right cervical lateral flexion motion was $7.64^\circ \pm 2.25^\circ$ ($P = .001$). Fourteen patients were diagnosed with intervertebral dysfunctions on the right side, whereas 11 patients showed cervical hypomobility on the left. Joint dysfunction at the C3 vertebra was the most prevalent (n = 16), followed by the dysfunction at the C4 vertebra (n = 9). The intervertebral radiological motion at the hypomobile side (mean 19.1, SD 2.1 mm) was $3.44 \pm 1.9$ mm less than the intervertebral radiological motion at the contralateral side (mean 22.6, SD 2.5 mm) with $P = .002$.

Conclusions: The lateral gliding test for the cervical spine was as good as a radiological assessment for the diagnosis of intervertebral dysfunctions in the lower cervical spine. (J Manipulative Physiol Ther 2005;28:610-616)

Key Indexing Terms: Neck Pain; Cervical Vertebrae; Motion; Joint Dysfunction; Lateral Gliding Test

Mechanical neck pain (MNP) affects 45% to 70% of the general population at some time during their lives.1 Mechanical neck pain can be defined as a neck disorder characterized by generalized neck and/or shoulder pain attributed to mechanical dysfunctions of the cervical spine.2 The exact pathology is not clearly understood but has been purported to be related to various anatomical structures including uncovertebral or intervertebral joints, neural tissues, disks, muscular disorders, and ligaments.3 Cervical joint dysfunction, known as somatic dysfunction, intervertebral joint dysfunction, chiropractic subluxation, or hypomobility by the various manipulating professions,4 is defined as a temporary reduction of mobility of a cervical segment.5

Spinal manipulative therapy is used by clinicians to reduce pain from intervertebral dysfunctions and to restore the biomechanical behavior of the spine.5 Clinical experience suggests that only zygapophyseal joints that are found to be hypomobile should be considered as candidates for high-velocity low-amplitude (HVLA) techniques. Therefore, manual diagnosis of joint hypomobility (ie, intervertebral dysfunction) constitutes the basis for deciding where to apply the HVLA technique. Intervertebral joint dysfunctions are usually diagnosed by physical examination. There are different clinical diagnostic tests aimed at assessing the passive intersegmental motion of the zygapophyseal joints; however, many of these tests lack scientific evidence to support their clinical relevance. However, Jull et al5 reported
that assessment of palpation, based on the guideline described by Maitland,\textsuperscript{6} could identify the presence and location of painful zygapophyseal joints with 100% sensitivity and specificity compared with diagnostic nerve blocks.

In clinical practice, two types of tests can be differentiated: one aimed at assessing the general mobility of the cervical spine (flexion, extension, rotation, lateral flexion) and others aimed at assessing the passive intersegmental motion of the zygapophyseal joints (springiness test, lateral gliding test). Studies analyzing the interexaminer reliability of tests assessing the general mobility of the cervical spine have reported $\kappa$ values ranging from 0.17 to 0.61.\textsuperscript{8,9} On the other hand, clinical tests assessing the passive intervertebral motion obtained $\kappa$ values ranging from 0.01 to 0.8 depending on the study and the clinical test.\textsuperscript{8-10} The substantial heterogeneity of the clinical tests studied in these trials makes it difficult to draw definite conclusions about their reliability. Moreover, we must take into account that upper cervical dysfunctions are assessed with different intervertebral motion tests than lower cervical dysfunctions. An old paper found that significant reliability in static and motion palpation was only found in lower cervical spine assessment.\textsuperscript{11}

Healthcare professions (eg, chiropractic, manual therapy, physical therapy) are in need of valid, consistent, and objective clinical tests for diagnosis of intervertebral dysfunctions. One of the most commonly used tests in clinical practice for diagnosis of cervical joint dysfunctions is the lateral gliding test.\textsuperscript{12} The aim of this test is to assess the passive lateral gliding and end-feel of each cervical vertebra. Therefore, a legitimate question is raised: is it possible to diagnose an intervertebral joint dysfunction in the mid and lower cervical spine with the lateral gliding test?

The aim of this study was to determine if the lateral gliding test for the cervical spine is a valid test, compared with radiological assessment, as tool for the manual diagnosis of intervertebral joint dysfunctions in patients presenting with MNP. The following hypothesis of this study was tested: findings on the lateral gliding test for cervical spine intervertebral joint dysfunction will show good agreement with radiological measurement. It was hypothesized that the intervertebral radiological motion, that is, the distance between the transverse process of the hypomobile vertebra and the transverse process of the adjacent vertebra, ipsilateral to the joint dysfunction would be reduced vs the intervertebral radiological motion at the contralateral side.

**METHODS**

**Subjects**

Twenty-nine patients, 13 males and 16 females, aged 20 to 44 years (mean 29.6, SD 6.7 years) presenting with MNP and referred by their primary care physician to a private physical therapy clinic in Madrid, Spain, from June to July 2003 were recruited to participate in this study. For the purpose of this study, MNP was defined as generalized neck and/or shoulder pain with mechanical characteristics including the following: symptoms provoked by maintained neck postures or by neck movement and/or by palpation of the cervical muscles. The health situation of the patients was clinically stable without current symptoms of any other concomitant chronic disease. All patients signed the informed consent form before beginning the trial. This study was supervised by the Department of Physical Therapy, Occupational Therapy, Physical Medicine and Rehabilitation of the Universidad Rey Juan Carlos, associated to the International School of Osteopathy in Madrid. It was approved by the Ethical Committee in Clinical Research of the university.

Inclusion requirements for patients to be participants were the following: MNP for at least 1 month with a negative extension-rotation test\textsuperscript{13}; asymmetry of 5\degree or greater between left and right cervical lateral flexion motions; lower cervical spine joint dysfunction diagnosed by the lateral gliding test; and at least 18 years old. Patients were excluded if they had any of the following: a history of neck trauma; diagnosis of fibromyalgia syndrome\textsuperscript{14}; history of cervical spine surgery; degenerative cervical alteration; diagnosis of cervical radiculopathy or myelopathy; and articular instability (sprain, fracture, luxation).

**Procedures**

All patients were first examined by therapist 1 for active cervical lateral flexion motion. For this assessment, subjects were asked to move symmetrically to the right cervical lateral flexion and then to the left, pausing at maximum end-range for an instant for the recorder to read and write the results of
measurement. Patients with at least 5° of asymmetry between both sides were included in the lateral gliding test exploration.

Later, therapist 2, who had more than 5 years of experience assessing joint dysfunctions and who was blinded to the goniometric assessment, examined the cervical spine for the presence of intervertebral joint dysfunction. Intervertebral dysfunctions were diagnosed by means of the lateral gliding test for the cervical spine as described by Greenman.12 For this test, the patient is supine with the cervical spine in a neutral position. The therapist places the fingers over the zygapophyseal joints of a specific cervical vertebra. The examining therapist laterally glides each vertebra from right to left and from left to right. Passive lateral gliding, end-feel, and quality of resistance were assessed. Once a restriction was identified in neutral, lateral gliding was performed in the same fashion with cervical spine flexion and cervical spine slight extension (Fig 1). The interpretation of this assessment is as follows: if the therapist identified restricted lateral gliding from right to left, a cervical hypomobility on the left side is recorded. Conversely, a lateral gliding restriction from left to right would indicate cervical hypomobility on the right side. An extension dysfunction is recognized by exaggerated restriction of lateral gliding of the cervical spine in the flexion position, but that restriction is reduced with the cervical spine in a position of extension. In this study, only extension dysfunctions were considered as they have been more frequently diagnosed in clinical practice.12

Once joint examination was complete, therapist 3, who had more than 4 years of experience in radiographic imaging and who was blinded to the aim of the study, performed the radiological examination. Each subject was seated upright in a chair in an x-ray suite. Patients were instructed to depress their shoulders to allow for clear visualization of the cervical spine. Patients were instructed to laterally flex their cervical spine to the right to end-range and then pause for an instant while an anterior-posterior cervical spine radiograph was taken. This procedure was then repeated with the patient in cervical lateral flexion to the left. A total of 50

### Table 1. Basic clinical data of each patient

<table>
<thead>
<tr>
<th>Patient</th>
<th>Age (y)</th>
<th>Sex</th>
<th>Right cervical lateral flexion</th>
<th>Left cervical lateral flexion</th>
<th>Intervertebral dysfunction and side</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>29</td>
<td>F</td>
<td>40°</td>
<td>50°</td>
<td>C3 left</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>M</td>
<td>40°</td>
<td>30°</td>
<td>C3 right</td>
</tr>
<tr>
<td>3</td>
<td>32</td>
<td>M</td>
<td>35°</td>
<td>40°</td>
<td>C3 left</td>
</tr>
<tr>
<td>4</td>
<td>26</td>
<td>M</td>
<td>25°</td>
<td>35°</td>
<td>C3 left</td>
</tr>
<tr>
<td>5</td>
<td>24</td>
<td>F</td>
<td>40°</td>
<td>31°</td>
<td>C4 right</td>
</tr>
<tr>
<td>6</td>
<td>21</td>
<td>F</td>
<td>45°</td>
<td>40°</td>
<td>C3 right</td>
</tr>
<tr>
<td>7</td>
<td>22</td>
<td>M</td>
<td>40°</td>
<td>35°</td>
<td>C3 right</td>
</tr>
<tr>
<td>8</td>
<td>39</td>
<td>F</td>
<td>40°</td>
<td>28°</td>
<td>C3 right</td>
</tr>
<tr>
<td>9</td>
<td>25</td>
<td>F</td>
<td>30°</td>
<td>40°</td>
<td>C3 left</td>
</tr>
<tr>
<td>10</td>
<td>28</td>
<td>M</td>
<td>40°</td>
<td>32°</td>
<td>C4 right</td>
</tr>
<tr>
<td>11</td>
<td>26</td>
<td>F</td>
<td>30°</td>
<td>40°</td>
<td>C4 left</td>
</tr>
<tr>
<td>12</td>
<td>30</td>
<td>M</td>
<td>30°</td>
<td>35°</td>
<td>C3 left</td>
</tr>
<tr>
<td>13</td>
<td>33</td>
<td>F</td>
<td>33°</td>
<td>40°</td>
<td>C4 left</td>
</tr>
<tr>
<td>14</td>
<td>25</td>
<td>F</td>
<td>35°</td>
<td>30°</td>
<td>C3 right</td>
</tr>
<tr>
<td>15</td>
<td>21</td>
<td>M</td>
<td>40°</td>
<td>35°</td>
<td>C3 right</td>
</tr>
<tr>
<td>16</td>
<td>44</td>
<td>F</td>
<td>40°</td>
<td>45°</td>
<td>C4 left</td>
</tr>
<tr>
<td>17</td>
<td>41</td>
<td>F</td>
<td>40°</td>
<td>30°</td>
<td>C3 right</td>
</tr>
<tr>
<td>18</td>
<td>23</td>
<td>M</td>
<td>40°</td>
<td>34°</td>
<td>C3 right</td>
</tr>
<tr>
<td>19</td>
<td>30</td>
<td>F</td>
<td>25°</td>
<td>30°</td>
<td>C4 left</td>
</tr>
<tr>
<td>20</td>
<td>35</td>
<td>M</td>
<td>37°</td>
<td>30°</td>
<td>C4 right</td>
</tr>
<tr>
<td>21</td>
<td>40</td>
<td>F</td>
<td>30°</td>
<td>39°</td>
<td>C3 left</td>
</tr>
<tr>
<td>22</td>
<td>30</td>
<td>M</td>
<td>40°</td>
<td>32°</td>
<td>C4 right</td>
</tr>
<tr>
<td>23</td>
<td>33</td>
<td>F</td>
<td>28°</td>
<td>35°</td>
<td>C3 left</td>
</tr>
<tr>
<td>24</td>
<td>28</td>
<td>F</td>
<td>40°</td>
<td>30°</td>
<td>C3 right</td>
</tr>
<tr>
<td>25</td>
<td>36</td>
<td>F</td>
<td>33°</td>
<td>25°</td>
<td>C4 right</td>
</tr>
</tbody>
</table>
anterior-posterior cervical spine radiographs were taken (two per patient).

Radiological Analysis of Intervertebral Motion

The radiograph analysis of intervertebral motion was performed by the third therapist. Markings were made on the radiographs to ascertain the distance, measured in millimeters, between the transverse process of the vertebrae making up the inferior joint surface of the hypomobile segment and the vertebrae making up the superior joint surface of the subjacent vertebrae. Tips of both transverse processes of each cervical vertebra were plotted on the radiographs. Both tips of the transverse process of the clinically identified hypomobile vertebra, determined by the lateral gliding test for the cervical spine, were connected with a continuous line. The same procedure was performed at the subjacent, that is, inferior vertebra. The distance, measured in millimeters, between the transverse process of the hypomobile vertebra and the transverse process of the subjacent vertebra was measured. This measurement was considered as the intervertebral motion at the hypomobile level (Fig 2). Because the marked pencil could influence radiographic measurements, intraobserver reliability was assessed (CCI = 0.92).

We obtained 25 radiographs in left lateral flexion and another 25 in right lateral flexion. One of these lateral flexions was considered ipsilateral to the side of the intervertebral dysfunction, and the other one was considered contralateral. The side of the dysfunction depended on the manual diagnosis through the lateral gliding test. The intervertebral radiological motion at the hypomobile zygapophyseal joint was measured from the radiograph with the patient in contralateral side flexion. Then, the intervertebral radiological motion at the zygapophyseal joint on the contralateral side was also measured to determine if the amount of motion between the hypomobile side and the contralateral side was different.

Instrumentation

A cervical goniometric device manufactured by Performance Attainment Associates (St Paul, Minn) was used for active cervical side flexion assessment. Lateral flexion score

<table>
<thead>
<tr>
<th>Patient</th>
<th>Intervertebral dysfunction</th>
<th>Goniometric cervical lateral flexion</th>
<th>Radiological intervertebral motion at the dysfunctional segment (mm)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>C3 Left</td>
<td>50°/40°</td>
<td>25/16</td>
</tr>
<tr>
<td>2</td>
<td>C3 Right</td>
<td>40°/30°</td>
<td>20/18</td>
</tr>
<tr>
<td>3</td>
<td>C3 Left</td>
<td>40°/35°</td>
<td>25/20</td>
</tr>
<tr>
<td>4</td>
<td>C3 Right</td>
<td>35°/25°</td>
<td>24/22</td>
</tr>
<tr>
<td>5</td>
<td>C4 Right</td>
<td>40°/31°</td>
<td>23/20</td>
</tr>
<tr>
<td>6</td>
<td>C3 Right</td>
<td>45°/40°</td>
<td>20/17</td>
</tr>
<tr>
<td>7</td>
<td>C3 Right</td>
<td>40°/35°</td>
<td>23/18</td>
</tr>
<tr>
<td>8</td>
<td>C3 Right</td>
<td>40°/28°</td>
<td>20/17</td>
</tr>
<tr>
<td>9</td>
<td>C3 Left</td>
<td>40°/30°</td>
<td>22/17</td>
</tr>
<tr>
<td>10</td>
<td>C4 Right</td>
<td>40°/32°</td>
<td>22/19</td>
</tr>
<tr>
<td>11</td>
<td>C4 Left</td>
<td>40°/30°</td>
<td>21/18</td>
</tr>
<tr>
<td>12</td>
<td>C3 Left</td>
<td>35°/30°</td>
<td>25/23</td>
</tr>
<tr>
<td>13</td>
<td>C4 Left</td>
<td>40°/33°</td>
<td>20/18</td>
</tr>
<tr>
<td>14</td>
<td>C3 Right</td>
<td>35°/30°</td>
<td>23/19</td>
</tr>
<tr>
<td>15</td>
<td>C3 Right</td>
<td>40°/35°</td>
<td>24/22</td>
</tr>
<tr>
<td>16</td>
<td>C4 Left</td>
<td>45°/40°</td>
<td>20/17</td>
</tr>
<tr>
<td>17</td>
<td>C3 Right</td>
<td>40°/30°</td>
<td>30/25</td>
</tr>
<tr>
<td>18</td>
<td>C3 Right</td>
<td>40°/34°</td>
<td>20/18</td>
</tr>
<tr>
<td>19</td>
<td>C4 Left</td>
<td>30°/25°</td>
<td>23/21</td>
</tr>
<tr>
<td>20</td>
<td>C4 Right</td>
<td>37°/30°</td>
<td>22/19</td>
</tr>
<tr>
<td>21</td>
<td>C3 Left</td>
<td>39°/30°</td>
<td>18/15</td>
</tr>
<tr>
<td>22</td>
<td>C4 Right</td>
<td>40°/32°</td>
<td>25/20</td>
</tr>
<tr>
<td>23</td>
<td>C3 Left</td>
<td>35°/28°</td>
<td>22/20</td>
</tr>
<tr>
<td>24</td>
<td>C3 Right</td>
<td>40°/30°</td>
<td>23/21</td>
</tr>
<tr>
<td>25</td>
<td>C4 Right</td>
<td>33°/25°</td>
<td>25/23</td>
</tr>
</tbody>
</table>

Mean 39.16°, SD 4.03°, P = .001 (based on 2-tailed Wilcoxon signed rank test)

Mean 31.25°, SD 4.25°, P = .002 (based on 2-tailed Wilcoxon signed rank test)

* Readers should take into account that the intervertebral radiological motion at the hypomobile side was measured from the radiograph with the patient in contralateral side flexion.
Lateral Gliding Test

Fernández-de-las-Pen˜as et al

RESULTS

Statistical Analysis

Data were analyzed with the SPSS package version 11.5 (SPSS, Chicago, Ill). Descriptive data were collected on all patients and then the group mean was calculated. The level and the side of the identified cervical joint dysfunction were recorded for each patient. Differences in the motion between the hypomobile and the contralateral side were analyzed with the nonparametric 2-tailed Wilcoxon signed rank test. Statistical analysis was conducted at a 95% confidence level because the testing was nonparametric and a small sample size was used. \( P < \alpha/2 = .025 \) was considered as statistically significant.

RESULTS

One patient was excluded because of the presence of only \( 3^\circ \) of asymmetry between left and right cervical lateral flexion; two patients were excluded because they were diagnosed with intervertebral joint dysfunctions at the Cl-C2 level; and the remaining one was excluded because of the presence of a flexion intervertebral dysfunction. Therefore, a total of 25 patients with MNP were included in the final study. The duration of neck symptoms of the included patients ranged from 1 to 6 months (mean 4.1, SD 1.2). The goniometrical assessment showed an asymmetry between left and right cervical lateral flexion motion of \( 7.64^\circ \pm 2.25^\circ \) (\( P = .001 \)). Fourteen patients were diagnosed with an intervertebral joint dysfunction on the right side, whereas 11 patients showed cervical hypomobility on the left side. Intervertebral dysfunction at the C3 vertebra was the most prevalent (\( n = 16 \) subjects, 64%), followed by hypomobility of the C4 vertebra (\( n = 9 \) subjects, 36%). The basic clinical data of each patient are summarized in Table 1.

Fifty anterior-posterior cervical spine x-rays were performed at maximum end-range of left (\( n = 25 \)) and right cervical lateral flexion (\( n = 25 \)). Radiological analysis was according to the side of the intervertebral joint dysfunction. Therefore, the intervertebral radiological motion at the hypomobile zygapophyseal joint was measured from the radiograph with the patient in contralateral side flexion, whereas the intervertebral motion at the contralateral hypomobile side was assessed from radiographs with the patient in side flexion homolateral to the cervical joint dysfunction (Fig 2). The intervertebral radiological motion at the hypomobile side (\( 19.1 \pm 2.1 \) mm) was less than the intervertebral radiological motion at the contralateral side (\( 22.6 \pm 2.5 \) mm). Therefore, the intervertebral radiological motion at the hypomobile side was \( 3.44 \pm 1.9 \) mm less than the contralateral hypomobile side (\( P = .002 \)). Table 2 summarizes the details of this analysis on each patient.

The results of patient 21 serve as an example (Table 2). Patient 21 was diagnosed with an intervertebral joint dysfunction on the left side at the C3 vertebra by manual diagnosis. Then, two anterior-posterior x-rays were performed. As this patient was diagnosed with an intervertebral joint dysfunction on the left side, right lateral flexion was considered as contralateral and left lateral flexion was considered as homolateral to the side of the hypomobility. Therefore, the intervertebral radiological motion between C3, the hypomobile vertebra, and C4, the subjacent vertebra, was assessed. Intervertebral motion at the hypomobile side was measured in right cervical lateral flexion, that is, side flexion contralateral to the side of the zygapophyseal joint dysfunction, and intervertebral motion at the contralateral hypomobile side was assessed in left cervical lateral flexion, that is, side flexion homolateral to the side of the zygapophyseal joint
dysfunction (Fig 4). In such a case, patient 21 showed an intervertebral radiological motion at the hypomobile side of 15 mm and an intervertebral radiological motion at the contralateral side of 18 mm, so the intervertebral radiological motion at the hypomobile side was 3 mm less than at the contralateral side.

**DISCUSSION**

There are many clinical tests used to diagnose intervertebral joint dysfunctions. One of the most common tests used in clinical practice for diagnosis of cervical joint dysfunctions is the lateral gliding test. We were unable to identify any study investigating the validity of this clinical test as a tool for the diagnosis of intervertebral joint dysfunctions in the lower cervical spine. Our study is the first to provide some evidence that the lateral gliding test for the cervical spine was as comparable with radiological assessment for the diagnosis of hypomobility in the lower cervical spine. Previous papers have analyzed the interexaminer reliability of different clinical tests aimed at assessing the passive intersegmental motion of the zygapophyseal joints. However, before assessing the reliability of a clinical test, it is necessary to validate it. Therefore, in clinical practice, a legitimate question can be raised before using any clinical test for manual diagnosis of joint dysfunction: is it possible to diagnose a zygapophyseal joint dysfunction with this clinical test? Is it true that the zygapophyseal joint that the therapist feels as hypomobile has a reduction of mobility? These questions should be answered before using any clinical test. The application of HVLA techniques is contingent upon the diagnosis of hypomobility (palpation, end-feel, gliding motion, etc). Therefore, the development, standardization, and validation of different manual methods to diagnose musculoskeletal pathologies are essential for our profession.

It is purported that joint hypomobility provokes a temporary reduction of mobility of a spinal segment. Therefore, it might be assumed that the hypomobile side will show less intervertebral motion than the contralateral side. Our study shows that the intervertebral radiological motion at the hypomobile side was 3.44 ± 1.9 mm less than the intervertebral radiological motion at the contralateral side (P = .002). This supports that the lateral gliding test might be comparable with radiological assessment of intervertebral motion restriction.

We have to take into account some aspects referring to manual diagnosis. In the lateral gliding test, as well as in other clinical tests, the passive lateral gliding, end-feel, and quality of resistance are considered in the diagnosis. Abnormal end-feel would be defined as the sensation that one would expect when the range of motion of a zygapophyseal joint is restricted. The capacity to discriminate this difference depends on extensive training and experience of the assessor. Moreover, end-feel could be different to different chiropractors. In addition, abnormal quality of resistance means that a greater than normal force has to be applied at the zygapophyseal joint to achieve the same degree of motion or gliding. As with abnormal end-feels, there are a variety of factors that might contribute to the quality of resistance: increase in joint viscosity, intra-articular adhesion, erosion of articular cartilage, and muscular spasm. Training and experience of the assessor play an important role in the diagnosis. These characteristics of manual diagnosis may be one of the reasons why interexaminer reliability of these tests had shown to be “poor to moderate” in previous studies. An interexaminer reliability poorer than expected can be caused by the fact that every therapist has their own individual methods for assessing intervertebral dysfunctions.

Jull et al reported that pathognomonic signs of symptomatic cervical zygapophyseal joint dysfunctions are abnormal end-feel, abnormal quality of resistance to motion, and reproduction of the patients’ pain. Mennel reported that relief of neck pain in patients with cervical spine involvement after spinal manipulative therapy is helpful in the diagnosis of intervertebral dysfunctions. In the present study, reproduction of the patients’ pain was not assessed because the aim of the study was to validate a mobility test. Moreover, in various recent papers, Jull et al reported that intervertebral dysfunctions can be determined validly without the subject providing pain-related feedback. Therefore, studies should also continue to test the concordance between results of manual examination and those from suitable medical diagnostic methods to further substantiate the contribution of manual diagnosis of intervertebral joint dysfunctions.

One limitation of the current study was the radiological assessment. This is the first study investigating the intervertebral radiological motion using radiological studies during side flexion motions. Yeomans analyzed the intervertebral radiological motion using radiological studies in flexion-extension motion. However, we must take into account that cervical hypomobilities are usually diagnosed either on the left or the right zygapophyseal joints; then, we would suggest that the radiological analysis of the intersegmental motion on that spinal segment should be a unilateral analysis, such as the analysis used in the present study.

To further investigate if the difference on the intervertebral radiological motion in the hypomobile side was located in that segment or if it might be found in other non-hypomobile segments, we also assessed the radiological intervertebral motion in all lower cervical segments. The analysis of this data did not show a statistical significant difference (based on nonparametric 2-tailed Wilcoxon signed rank test) between both sides (P > .2), obtaining radiological differences ranging from 0.8 to 1.1 mm depending on the cervical segment. Therefore, we might
conclude that restricted intersegmental motion seen in the hypomobile segment was not caused by the radiological assessment, and that the lateral gliding test permitted the manual diagnosis of that restricted zygapophysial motion.

Previous papers\textsuperscript{21,22} have documented an intraobserver inherent error of 0.6 mm (SD \pm 0.8) and 0.4 mm (SD \pm 0.1) using stress radiography in flexion-extension motion in the cervical spine. Because we were unable to locate any study analyzing the reliability of stress radiography in lateral flexion motion, we assessed the intraobserver reliability of our radiographies (CCI = 0.92), obtaining an inherent error ranging from of 0.3 to 0.6 mm depending on the cervical segment measured. Although the obtained error was similar to that reported in the aforementioned studies, the reliability of this technique requires further investigation. In future studies, more sophisticated techniques, such us the cineradiography,\textsuperscript{23} to assess the intersegmental motion of the cervical spine should be considered.

The small sample size is another limitation of this study. Type II errors could have happened; therefore, it is recommended to repeat the same study with a larger sample size and with healthy control subjects. However, as we used a nonparametric test in the statistical analysis, the sample size was considered sufficient for the aim of the study. Finally, we have to recognize that the goniometric assessment might not have been a very reliable measurement tool. The inclusion of an asymmetric motion on cervical lateral flexion motion permitted us more restrictive inclusion criteria. However, we think that the assessment of cervical motion by cervical goniometry does not alter the results of the study as our aim was to compare the results between manual diagnosis and radiological assessment.

CONCLUSIONS

The lateral gliding test for the cervical spine was as good as a radiological assessment for the diagnosis of intervertebral joint dysfunction in the lower cervical spine in this group of patients. These results support that the lateral gliding test for the cervical spine may be a valid manual tool for the diagnosis of restricted intervertebral mobility in the lower cervical spine.

REFERENCES

LITERATURE REVIEWS

AN ANALYSIS OF THE ETIOLOGY OF CERVICAL ARTERY DISSECTIONS: 1994 TO 2003

Michael T. Haneline, DC, MPH and Gary N. Lewkovich, DC

ABSTRACT

Objective: To provide a literature review of the etiologic breakdown of cervical artery dissections.

Methods: A literature search of the MEDLINE database was conducted for English-language articles published from 1994 to 2003 using the search terms cervical artery dissection (CAD), vertebral artery dissection, and internal carotid artery dissection. Articles were selected for inclusion only if they incorporated a minimum of 5 case reports of CAD and contained sufficient information to ascertain a plausible etiology.

Results: One thousand fourteen citations were identified; 20 met the selection criteria. There were 606 CAD cases reported in these studies; 321 (54%) were internal carotid artery dissection and 253 (46%) were vertebral artery dissection, not including cases with both. Three hundred seventy-one (61%) were classified as spontaneous, 178 (30%) were associated with trauma/trivial trauma, and 53 (9%) were associated with cervical spinal manipulation. If one apparently biased study is dropped from the data pool, the percentage of CADs related to cervical spinal manipulation drops to approximately 6%.

Conclusions: The case series that were reviewed in this article indicated that most CADs reported in the previous decade were spontaneous but that some were associated with trauma/trivial trauma, and a minority with cervical spine manipulation. This etiologic breakdown of CAD does not differ significantly from what has been portrayed by most other authors. (J Manipulative Physiol Ther 2005;28:617-622)

Key Indexing Terms: Cervical Artery Dissection; Trauma; Manipulation; Neck; Vertebral Artery

Cervical artery dissection (CAD) is an uncommon condition that begins as a tear or defect of the intimal lining of one of the major arteries in the neck that supplies blood to the brain. Subsequent to this insult and because of arterial pressure, blood penetrates into the vessel wall. The intima will usually separate from the medial layer, typically in the direction of blood flow. Subintimal dissection may cause luminal narrowing or even occlusion leading to cerebral ischemia (Fig 1). Cervical artery dissection may also manifest as a subadventitial dissection, which may result in the formation of a pseudoaneurysm (Fig 2). At times, CAD occurs with no associated tear of the intima; instead, there is hemorrhage of the vasa vasorum, allowing the formation of an intramural hematoma (Fig 3).

The precise pathogenesis of CAD is not clear in most cases. Trauma, ranging from trivial to severe, is often implicated. Trivial traumas associated with CADs are numerous and include such common events as turning the head to back up a car, receiving cervical spinal manipulation, dancing, or having dental work. More commonly, however, CADs occur spontaneously, meaning that no apparent causative agent could be determined from the available data. This conclusion has prompted some authors to suggest that there must be an underlying arteropathy, which predisposes certain individuals to dissection.

Several authors have reported that the incidence of arterial dissection is higher among patients with fibromuscular dysplasia, Marfan’s syndrome, migraine, or hypertension. Smoking and the use of oral contraceptives are behaviors that may be associated with higher rates of dissection.

a Associate Professor, Palmer College of Chiropractic West, San Jose, Calif.
b Private Practice.
Sources of support: No external funding was provided for this study.
Submit requests for reprints to: Michael T. Haneline, DC, MPH, Palmer College of Chiropractic West, 90 E. Tasman Drive, San Jose, CA 95134 (e-mail: michael.haneline@palmer.edu).
0161-4754/$30.00
Copyright © 2005 by National University of Health Sciences.

617
infection and mild hyperhomocysteinemia have also been identified as risk factors and could be triggers for spontaneous CAD.10,11

The distinction between spontaneous dissection and dissection due to trivial trauma is ambiguous. The clinical presentation of these two entities does not differ significantly, making it difficult to determine whether a dissection is spontaneous or secondary to trivial trauma.12,13 Indeed, if minimal “trauma” such as bowling and coughing is capable of precipitating CADs, then no dissection may truly be considered spontaneous. As a result, some authors have combined trivial trauma etiologies with spontaneous dissections when reporting cases in the literature.

A report highlighted a case of internal carotid artery dissection (ICAD) that was associated with French horn playing. The authors pointed out that there were concomitant aberrations of the patient’s connective tissue and mildly elevated plasma homocysteine levels that may have predisposed the patient to dissection. They stated that the growing number of anecdotal reports linking everyday physical activities to the development of CADs may overestimate the pathogenic role of trauma.14

The incidence of vertebral artery dissection (VAD) has been estimated by Schievink et al16 to be 1 to 1.5 per 100,000 population yearly.15 Two epidemiologic studies have estimated the incidence of spontaneous ICAD as 2.6 and 2.9 cases per 100,000 per year.16,17 When comparing the two types of dissections, ICAD has been reported to occur 3 to 5 times more frequently than VAD.18,19 It has been reported that there are more than 7000 cases per year of ICADs alone in the United States.20

Headache and neck pain occur in 50% to 80% of CADs, and may be the only warning symptoms of impending dissection.9,19 This is especially true when the headaches are described as new or unusual and characterized as sharp. Headaches associated with CADs often start suddenly and, most frequently, are unilateral on the side of arterial dissection. Furthermore, a CAD is even more likely when signs and symptoms of cerebral ischemia, Horner’s syndrome, or pulsatile tinnitus accompany the head and/or neck pain.

Because the earliest symptoms of a CAD are often neck pain and/or headache, it is likely that a victim of this disorder will seek assistance from one of several types of healthcare providers. Medical doctors, acupuncturists, naturopaths, herbologists, and chiropractors are all possible sources of attempted relief. Given the thousands of spontaneous CAD cases that occur each year in the United States alone, the likelihood that chiropractic physicians will encounter patients with CAD is relatively high.

There have been numerous reports in the medical literature regarding the proportion of CADs thought to be spontaneous, related to trauma/trivial trauma, or related to cervical manipulation.21-25 The proportion of CADs that can be attributed to manipulation, in particular, is uncertain.
Accordingly, the purpose of this article is to produce a reasonable estimate of the etiologic breakdown based upon a literature search of suitable studies published during the previous decade.

METHODS

A literature search was performed to locate case series that dealt with etiologic factors of CAD during the preceding 10-year period. The MEDLINE database was searched for English-language articles published during the period extending from 1994 to 2003. Search terms included cervical artery dissection, vertebral artery dissection, and internal carotid artery dissection. In addition, a cross-tabulation was performed from the reference section of the selected articles to find other suitable studies. The inclusion criteria for article selection encompassed all studies that reported 5 or more cases of CAD and also provided enough information to ascertain the supposed etiology of the dissections.

In an attempt to minimize possible selection bias, only studies that included 5 or more cases were included in this research. Single case reports often have dealt with an atypical patient whose condition was highlighted to educate the reader about an unusual medical situation. Authors typically select rare or unusual conditions as topics for case reports. Where CAD is involved, certain types of trauma or manipulation-related cases may be over-represented because they tend to be more interesting than CADs with an uneventful or mundane history. This type of problem was

| Table 1. Cervical artery dissection case series reported from 1994 to 2003, where factors of etiology could be determined

<table>
<thead>
<tr>
<th>n</th>
<th>ICAD/VAD</th>
<th>Spontaneous</th>
<th>Trivial or other trauma</th>
<th>Manipulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bassi et al(^{27})</td>
<td>49</td>
<td>36/13</td>
<td>31</td>
<td>18</td>
</tr>
<tr>
<td>Chabrier et al(^{26})</td>
<td>5(^{b})</td>
<td>3/2</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Dziewas et al(^{1})</td>
<td>126</td>
<td>78/46(^{3})</td>
<td>77</td>
<td>29</td>
</tr>
<tr>
<td>Gonzales-Portillo et al(^{29})</td>
<td>27</td>
<td>18/9</td>
<td>22</td>
<td>3</td>
</tr>
<tr>
<td>Pezzini et al(^{30})</td>
<td>34</td>
<td>20/14</td>
<td>32</td>
<td>1</td>
</tr>
<tr>
<td>Camacho et al(^{11})</td>
<td>5</td>
<td>4/1</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Brandt et al(^{35})</td>
<td>65</td>
<td>29/7(^{c})</td>
<td>55</td>
<td>8</td>
</tr>
<tr>
<td>Norris et al(^{12})</td>
<td>74</td>
<td>21/53</td>
<td>14</td>
<td>39</td>
</tr>
<tr>
<td>Saeed et al(^{35})</td>
<td>26</td>
<td>8/26</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>Shin et al(^{13})</td>
<td>8(^{e})</td>
<td>8/8</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Ahmad et al(^{34})</td>
<td>18</td>
<td>10/8</td>
<td>2</td>
<td>14</td>
</tr>
<tr>
<td>Auer et al(^{15})</td>
<td>19</td>
<td>5/19</td>
<td>7</td>
<td>12</td>
</tr>
<tr>
<td>Brandt et al(^{36})</td>
<td>25</td>
<td>17/8</td>
<td>20</td>
<td>4</td>
</tr>
<tr>
<td>de Bray et al(^{17})</td>
<td>7</td>
<td>8/7</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Mascalchi et al(^{38})</td>
<td>14</td>
<td>5/14</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>Sidhu et al(^{39})</td>
<td>5</td>
<td>5/5</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Bakke et al(^{30})</td>
<td>14</td>
<td>12/2</td>
<td>12</td>
<td>2</td>
</tr>
<tr>
<td>Desfontaines and Despland(^{24})</td>
<td>60</td>
<td>60/(^{§})</td>
<td>50</td>
<td>10</td>
</tr>
<tr>
<td>Sturzenegger(^{41})</td>
<td>14</td>
<td>5/14</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>Schievink et al(^{7})</td>
<td>11(^{b})</td>
<td>8/2(^{c})</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Total</td>
<td>606</td>
<td>321/253</td>
<td>371</td>
<td>178</td>
</tr>
</tbody>
</table>

\(^{§}\) Not included in the study.

\(^{a}\) Selection criteria: cervical artery dissection case series of 5 or more cases reported between 1994 and 2003 (inclusive); sufficient information was presented regarding etiology; and cases were not selected based on type of etiology (ie, all manipulation-related or all spontaneous).

\(^{b}\) Children; does not include intracranial cases.

\(^{c}\) Does not include cases that experienced both ICAD and VAD.

\(^{d}\) Personal communication with T Brandt, MD.

\(^{e}\) Intracranial CAD cases were excluded.

| Table 2. Breakdown of CADs involving the ICA vs the VA (when reported) that were associated with manipulation from case series reported from 1994 to 2003

<table>
<thead>
<tr>
<th>Manipulation-related CAD</th>
<th>N</th>
<th>ICAD</th>
<th>VAD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dziewas et al(^{1})</td>
<td>19(^{a})</td>
<td>5</td>
<td>14</td>
</tr>
<tr>
<td>Gonzales-Portillo et al(^{29})</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Brandt et al(^{36})</td>
<td>2</td>
<td>1</td>
<td>1(^{b})</td>
</tr>
<tr>
<td>Saeed et al(^{35})</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Ahmad et al(^{34})</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Brandt et al(^{36})</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>de Bray et al(^{17})</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Mascalchi et al(^{38})</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Sturzenegger(^{41})</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>34</td>
<td>8</td>
<td>26</td>
</tr>
</tbody>
</table>

\(^{a}\) Not included in the study.

\(^{a}\) Does not include cases that experienced both ICAD and VAD.

\(^{b}\) Personal communication with T Brandt, MD.
apparent in a study by Frisoni and Anzola\textsuperscript{26} where 3 cases of CAD were presented in an article highlighting the alleged relationship of VAD with neck motion, when all 3 cases were manipulation-related.

**RESULTS**

A total of 1014 citations were identified, 20 of which reported sufficient numbers of CAD cases to be included in this analysis. There were 606 CAD cases reported among the 20 studies; 321 (54\%) were ICAD and 253 (46\%) were VAD. When only considering studies that included both types of dissection, there were 256 (61\%) ICADs and 165 (39\%) VADs. Most of the authors (58\%) incorporated both types of dissection in the same study, whereas 42\% exclusively reported either VAD or ICAD cases.

There were 371 (61\%) CADs listed as spontaneous, 178 (30\%) associated with trauma/trivial trauma, and 53 (9\%) associated with manipulation. Table 1 lists all of the case series that were included. Seven studies with sufficient numbers of CAD cases to meet the inclusion criteria did not differentiate spontaneous cases from other possible causes and were therefore excluded from the analysis.\textsuperscript{42-48} When specifically considering studies reporting manipulation-related CADs that made a distinction between the occurrence of dissection in the internal carotid artery (ICA) vs the vertebral artery (VA), there were 34 cases, 8 (24\%) involving the ICA and 26 (76\%), the VA (Table 2).

**DISCUSSION**

The breakdown of CAD etiologies that has been presented in this review provides a reasonable depiction of what is contained in the recently published literature. However, because there are no widely-agreed-upon definitions for the categorization of CAD etiologies, these figures may not accurately reflect the true breakdown. Categorizing CADs as spontaneous or trivial trauma—related may seem, at first glance, to be a straightforward task. Nevertheless, the issue of trauma/trivial trauma and its connection to CAD is controversial because a temporal link is generally the only evidence that is available when determining culpability. Brandt and Grond-Ginsbach\textsuperscript{49} pointed out that, based on clinical and histopathologic studies, mechanical trauma was not an “important and frequent cause” for the development of CAD. They indicated that although mechanical stretching or compression of the cervical arteries may be capable of triggering CAD, mechanical injury may not be the most important risk factor in such cases. Moreover, supposed “traumas” or “trivial traumas” may not have been involved at all in some instances. This concept was substantiated by a histopathologic study of 50 ICADs that found less than 10\% of the preparations showed any indication of mechanical damage to the artery.\textsuperscript{50}

Pathogenic arterial wall weakness is a predisposing factor that is often present in CAD patients and may play a more important causative role than mechanical factors.\textsuperscript{49} In cases where there are no adverse symptoms within a reasonably close temporal proximity to trauma/trivial trauma, there will always be doubt regarding its connection with any given CAD. Furthermore, patients with arterial disorders are plausibly at risk for developing CAD when subjected to virtually any sort of head or neck motion.

It was previously mentioned that ICADs have been reported to occur 3 to 5 times more frequently than VADs. However, this review points toward a ratio of approximately 3 ICADs for every two VADs. The reader should bear in mind that the accuracy of this ratio is dependent upon the various methods of data collection used in the collected research.

The relative proportions of CAD etiology among the studies included in this analysis were fairly consistent. One exception to this consistency does exist, however. The Norris et al\textsuperscript{12} case series presented data highly atypical of past studies, especially about the reported ratio of VAD to ICAD cases. All of the other studies showed a higher number of ICADs compared with VADs. In fact, the overall ratio of ICAD to VAD was 300:200, not including the Norris et al data. In contrast, the Norris et al ratio was 21:53, with VAD cases more than doubling the number of CAD cases. Another inconsistent finding in the Norris et al study concerned the authors’ statement that more than 80\% of their CAD cases were related to trivial trauma or manipulation, with 28\% being related to cervical manipulation alone. These figures were much higher than what was reported in any of the other studies reviewed.

In an attempt to appreciate how the above referenced discrepancies might have occurred, we offer a very brief critique of that Norris et al\textsuperscript{12} study. There appears to have been systematic methodological biases that were integrated into the study’s design. For example, the survey instrument used to gather the study’s CAD data, available from the Canadian Stroke Consortium (www.strokeconsortium.ca) contained a serious problem with respect to its poor construction. There was a leading question in a foundational portion of the questionnaire that may have caused the physician respondents to provide biased answers. The question dealt with whether trauma was related to the CAD under consideration and reads, “Cervical trauma/manipulation.” Because manipulation is only one of the many trivial events that are thought to be associated with CAD, it should not have been singled out and, thus, presented as a leading question. Bias was again evident when the word “manipulation” was used 13 times in this short article about the association of all sudden neck movement to CAD.

Reviewing all the studies included in this analysis, 30\% of the CADs were found to be associated with trauma/trivial
trauma and 9%, with manipulation. These percentages include the Norris et al.\textsuperscript{32} data, although they appeared to be biased against manipulation. If one excludes the suspect Norris et al. data, the percentage of CADs related to trauma/trivial trauma drops to 26%, and the percentage related to manipulation decreases to 6.1%.

Rothwell et al.\textsuperscript{47} recently carried out a population-based study which appears to be the most reliable investigation to date regarding the percentage of VADs associated with manipulation. Its reliability lies in the fact that the study included all VADs occurring in the Province of Ontario, Canada that were admitted to any acute-care facility during a 6-year period. This method minimized selection biases that have plagued most other studies. The figures they presented suggest that approximately 1.6% of persons with VAD received cervical manipulation in the week before the onset of definitive cerebral ischemic findings. This percentage was significantly less than 9%, which represented the average of all the studies included herein. Accordingly, the true proportion of VADs ostensibly associated with manipulation remains elusive, but it is most likely less than 9%.

**CONCLUSION**

This work provides an estimate of the etiologic breakdown of CADs reported during the previous decade. The case studies reviewed indicate that, during the specified period, most CADs were labeled as spontaneous, but some were associated with trauma/trivial trauma, and a minority was associated with cervical manipulation. The proportion of CAD etiologies presented does not differ significantly from what has been presented by other authors, with the exception of Norris et al.\textsuperscript{32}

The etiology of CAD has been difficult to investigate because of its relatively uncommon rate of occurrence, the delay that is thought to transpire between the time of dissection and onset of symptoms, and the lack of available histopathologic information in most cases. Moreover, the temporal relationship of an event with the onset of symptoms does not necessarily imply causation. As stated previously, the etiology of CAD is complex and often multifactorial, especially when trivial trauma and manipulation are involved. Other risk factors or risk conditions, such as fibromuscular dysplasia, Marfan’s syndrome, migraine, hypertension, smoking, use of oral contraceptives, recent infection, and mild hyperhomocysteinemia, should be considered in any given case.

**REFERENCES**


THE BIOMECHANICAL AND CLINICAL SIGNIFICANCE OF THE LUMBAR ERECTOR SPINAE FLEXION-RELAXATION PHENOMENON: A REVIEW OF LITERATURE

Christopher J. Colloca, DC, and Richard N. Hinrichs, PhD

ABSTRACT

Objectives: The aim of this study was to review the biomedical literature to ascertain the biomechanical and clinical significance of the lumbar erector spinae flexion-relaxation phenomenon (FRP).

Data Sources: Index Medicus via PubMed, the Noble Science Library’s e-journal archives, and the Manual Alternative and Natural Therapy Index System databases were searched using the same search terms.

Discussion: The presence of the FRP during trunk flexion represents myoelectric silence consistent with increased load sharing of the posterior discoligamentous passive structures. Passive contributions from erector spinae stretching during the flexion posture and active contributions from other muscles (quadratus lumborum and deep erector spinae among others) further assist in load sharing in the trunk flexion posture. A number of studies have shown differences in the FRP between patients with chronic low back pain and healthy individuals, and the reliability of the assessment. Persistent activation of the lumbar erector spinae musculature among patients with back pain may represent the body’s attempt to stabilize injured or diseased spinal structures via reflexogenic ligamentomuscular activation thereby protecting them from further injury and avoiding pain.

Conclusions: The myoelectric silencing of the erector spinae muscles in the trunk flexion posture is indicative of increased load sharing on passive structures, which tissues have been found to fail under excessive loading conditions and shown to be a source of low back pain. The studies that show differences in the presence of the FRP among patients and control subjects are encouraging for this type of clinical assessment and suggest that assessment of the FRP is a valuable objective clinical tool to aid in the diagnosis and treatment of patients with low back pain. (J Manipulative Physiol Ther 2005;28:623-631)

Key Indexing Terms: Biomechanics; Electromyography; Low Back Pain; Lumbar Vertebrae; Flexion-Relaxation Phenomenon; Trunk Flexion

Movements in the lumbar spine, including flexion and extension, are governed by a complex neuromuscular system involving both active (muscle) and passive (vertebral bones, intervertebral disks, ligaments, tendons, and fascia) components. Common among spinal disorders are disruption to the neuromuscular balance and load sharing of the spinal tissues, ultimately resulting in pain and disability, and an enormous economic burden to society.

In the assessment of patients with lumbar complaints, measuring the electromyographic (EMG) activity of the trunk musculature is one objective means used by biomechanists and clinicians to assess the function of the lumbar spine. The clinical utility of the use of electromyography, however, is controversial in the diagnosis of patients with low back pain without lower extremity symptoms.

There is evidence to suggest that EMG differences exist between patients with back pain and healthy subjects during dynamic flexion tasks performed at peak flexion. To this extent, several studies have examined the apparent myoelectric silencing of the low back extensor musculature during a standing to full trunk flexion maneuver or the flexion-relaxation phenomenon (FRP). The electrical signal reduction or silence that occurs in healthy subjects during lumbar spine flexion has been hypothesized to represent...
the extensor musculature being relieved of its moment-supporting role by the passive tissues, particularly the posterior spinal ligaments. Likewise, a failure of the muscles to relax in patients with back problems is indicative of heightened erector spinae resting potentials or underlying back muscle spasticity.

The FRP is modulated by a number of factors including the magnitude of applied load, loading rate, and patient clinical status. Creep developed during a short static lumbar flexion has also been found to elicit significant changes in the muscular activity pattern of the FRP. Although there are many factors to consider, understanding the biomechanical importance of the load sharing between the erector spinae musculature and passive lumbar spinal tissues will aid in the understanding of the FRP from both a biomechanical and clinical standpoint. Moreover, reviewing the biomechanical implications of the FRP and reported differences among patients with back pain and healthy subjects assists in understanding the benefits, limitations, and clinical utility of this type of EMG assessment. Thus, the purpose of this study is to review the biomedical literature to ascertain the biomechanical and clinical significance of the lumbar erector spinae FRP.

**METHODS**

A review of the biomedical literature on the lumbar erector spinae FRP was conducted by searching the National Library of Medicine’s *Index Medicus* via their PubMed database, the Noble Science Library’s e-journal archives, and the Manual Alternative and Natural Therapy Index System using the following search terms: *biomechanics, electromyography, erector spinae muscle, flexion-relaxation, lumbar spine, and trunk flexion*. Keywords were input in various combinations, always including the terms flexion-relaxation or trunk flexion, and relevant citations were noted. Collectively, these searches revealed 36 articles generally related to the research question in the current study, ultimately resulting in 16 specific articles on the FRP. These 16 articles among...
DISCUSSION

Physical activities involving full trunk flexion are common in activities of daily living, occupational demands, and sport. Thus, knowledge of the biomechanics and clinical implications of trunk flexion is important. Lumbar spinal tissue sprain and strain have been reported in lumbar flexion postures, and patients with low back disorders often avoid such positions. Understanding the transfer of tissue loads in the trunk flexion posture thus assists in the understanding of normal trunk biomechanics, mechanisms of injury, and the consequential pathogenesis of low back pain.

Several biomechanical and clinical studies have examined the apparent myoelectric silence of the lumbar spine extensor musculature during lumbar flexion from a neutral upright standing posture. The FRP refers to a pattern of muscle activity during trunk flexion in which the lumbar muscles ultimately relax at what appears to be a distinct point in the lumbar flexion range of motion (Fig 1). Floyd and Silver first described the term flexion-relaxation of the lumbar extensor musculature using EMG and suggested that the passive lumbar posterior elements, namely, the posterior spinal ligaments and intervertebral disks, supplied the needed moment during full flexion in the absence of erector spinae muscle activity. The mechanism for the silencing of the erector spinae muscles during trunk flexion has been proposed to result from stimulation of stretch receptors in the posterior discoligamentous tissues during the flexed posture, acting to reflexogenically inhibit erector spinae activity. A number of biomechanical studies have since examined the transfer of loads among tissues during lumbar flexion, and several clinical studies have further investigated the significance of the presence or absence of the FRP in patients with lower back pain (Fig 2). Such research is not only important to understand the biomechanical consequences of the trunk flexion posture but to further understand the clinical utility of using the FRP as an objective outcome measure to discriminate patients with low back pain. The FRP is an appealing quantitative test for adding objectivity to a movement in which pain inhibition and...
voluntary effort limitations may confound the examiner’s ability to assess actual lumbar flexibility.\textsuperscript{15}

In normal trunk flexion with the knees straight, the 5 lumbar vertebral segments flex forward during the first 50\degree \textendash 60\degree, followed by the pelvis rotating between the hips.\textsuperscript{16} At 75\% to 85\% of trunk flexion, the lumbar spine reaches its maximum range of motion, whereas the pelvis is providing terminal flexion to achieve the final stages of total trunk flexion.\textsuperscript{7} It has been postulated that, at this point, the passive noncontractile soft tissues (intervertebral disks, ligaments, fascia) are providing most of the spinal support, and little erector spinae activity is required to maintain this posture.\textsuperscript{6,7} Interestingly, lumbar flexion is the usual posture for stoop labor, a position selectively chosen by workers in many types of field labor for its efficiency (lack of voluntary muscle effort).\textsuperscript{17} From a clinical standpoint, patients with chronic low back pain (CLBP) have been reported not to achieve flexion-relaxation because of an abnormal neuromuscular coordination between the trunk and hip movements.\textsuperscript{9}

**Task Performance**

To assess the EMG activity of the lumbar trunk muscles during trunk flexion, use of a data acquisition system is necessary. Surface electromyographic electrodes are attached to the skin at the levels of T12 and L3 over the belly of the erector spinae muscles, and a trunk measurement device (electrogoniometer) is attached at the levels of T12 and S1. A reference ground electrode is shown on the right olecranon. Simultaneous erector spinae muscle EMG activity can be measured as a function of lumbar range of motion with this experimental setup.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig3}
\caption{Typical electrode and electrogoniometer placement for dynamic surface EMG assessment. Surface EMG electrodes are placed overlying the erector spinae musculature at the levels of T12 and L3, and an electrogoniometer (paddles) is placed at the level of T12 and S1. A reference ground electrode is shown on the right olecranon. Simultaneous erector spinae muscle EMG activity can be measured as a function of lumbar range of motion with this experimental setup.}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig4}
\caption{Lumbar flexion test performance. (A) The standing patient flexes forward (B) and returns to the standing position (A) during the flexion range of motion test. Surface electromyographic recordings are obtained from the erector spinae muscles, and simultaneous range of motion is monitored by means of a dual paddle electrogoniometer.}
\end{figure}

S1 to quantify the lumbar motion during the test (Fig 3). Alternatively, a video motion analysis system can be used and synchronized with the EMG signals obtained to assess the
EMG response in relation to the trunk motion. In the standing position, the patient is then asked to stand quietly in their neutral upright posture and bend forward to the point of full flexion at a speed that they feel comfortable with as EMG signals are recorded (Fig 4). Generally, 3 repetitions are performed to allow the examiner to observe any discontinuity of the task effort and to obtain mean values of range of motion and EMG activity. During the test, surface EMG activity is recorded in relation to the spinal motion during the trials and peak-peak amplitude of the EMG signals can be assessed. Several variables such as patient population, recording technique, data analysis and interpretation (defining on and off), normalization techniques, and task performance (including controlling for velocity of motion) among others are responsible for differences in results among studies investigating the FRP. In addition, the onset and cessation of the myoelectric silence of the FRP can be influenced by several factors including lumbar lordosis, general laxity of the joints, strength and relative length of the muscles of the trunk and hip, coordination of trunk and hip movements, and the velocity of the flexion extension movements.18

The notion that the FRP represents a dichotomous pattern of muscles being either on or off likely is misleading. Physiologically, lumbar muscles show electrical activity because they are active and the degree of electrical activity is roughly proportional to the rate and amplitude of muscle fiber firing. Nevertheless, defining EMG activity to denote when muscles are active is necessary to proportion and interpret EMG amplitude results from raw data that have been collected. Exploring the different assessment and analytical techniques presented in the literature serves to point out differences in results reported in the literature and subsequently understand subsequent limitations.

Quantifying the FRP and Load Transfer During Trunk Flexion

A number of studies have shown the presence of the FRP during upright trunk flexion in asymptomatic subjects.1,7,12,15,19-22 Consistent with these studies, recently, Solomonow and coworkers1 recorded EMG activity at the L3 through L4 level over the erector spinae musculature during upright trunk flexion tasks and quantified trunk joint angles using a two-dimensional video–based motion analysis system in 49 asymptomatic subjects. To nullify time shifts of the EMG data, a digital algorithm to yield a mean absolute value with a time shift of 100 milliseconds was performed and EMG-on and EMG-off were defined, with the latter representing EMG silence. Flexion-relaxation was observed in all subjects during trunk flexion, with the mean EMG-off lumbar flexion angle for men and women ranging between 46° to 50° of trunk flexion. Studies showing the FRP in asymptomatic subjects will be further reviewed hereinafter, in context with their biomechanical and clinical significance.

In any static position of trunk flexion, the hip flexion is maintained by balance of the torque of the upper body weight resisted by a combination of the tension and mass of the structures posterior to its axis, and vertebral flexion is maintained by the force of the upper body weight resisted by combination of tension and compression of the vertebral structures.7 Thus, it is rationalized that the FRP of the erector spinae musculature in the trunk flexion posture represents a “switching off” of electrical activity due to equilibrium being achieved between the torques because of gravity and the extension torque provided by the stretched posterior vertebral elements.18 The erector spinae thus plays the key role of balancing the two passive forces and providing controlled movement of vertebral flexion through its eccentric action.

To understand the biomechanical consequences of the FRP, other research has been conducted to determine the load sharing among lumbar spinal tissues in the flexed posture. Schultz et al7 measured trunk muscle EMG in 8 healthy subjects during upright lumbar flexion with various loading conditions, and biomechanical analyses were used to estimate the subsequent loads imposed on the lumbar trunk structures. In these analyses, the net support reaction needed for equilibrium across a transverse section at the level of L3 was computed followed by analysis of muscle action that could supply the net reaction derived using an optimization technique. The resulting spine compression and shear loads were then estimated as well. The authors found that, during lumbar flexion, myoelectric signals of the erector spinae muscles were substantially smaller compared with quiet upright standing despite the need to develop posterior tissue tensions equivalent to erector spinae contractions exceeding 700 N on each side at 40° or more of flexion to maintain equilibrium. Based upon the data obtained and the biomechanical analyses performed in this study, it was clear that passive tissue forces are required to sustain the load requirements in the trunk flexion posture.

The resistance to the large flexion moments generated at the trunk has been further investigated by McGill and Kippers6 who examined the loads on individual tissues during the performance of the FRP using an anatomically detailed model of the lumbar spine using vertebral displacement and myoelectric signals to estimate individual muscle and passive tissue force-time histories. The authors reported that, in full trunk flexion, loading of the interspinous and supraspinous ligaments and posterior annulus of the intervertebral disk in particular were found to be high relative to their failure tolerances. Flexion of the trunk places the posterior passive elements of the spine at risk for failure, which is consistent with injury mechanisms in many patients with low back pain. McGill and Kippers6 further noted that although many muscles show force generation consistent with electrical activity from neural activation, forces are still predicted in the lumbar extensor muscles because of passive stretching far beyond their resting length in the trunk-flexed posture, enabling the muscles to generate forces in the absence of neural activation. Thus, this was the first study to propose that although the erector spinae muscles become
electrically silent during trunk flexion, they still appear to be generating force elastically through passive stretching.

Dolan et al.\textsuperscript{23} further investigated the passive tissue contributions of the spine during lifting in the trunk flexion posture in 149 healthy subjects. In their study, subjects in the trunk-flexed posture pulled upward with steadily increasing force on a floor-mounted load cell, whereas EMG activity was recorded from the erector spinae muscles at L3 and T10. Extensor moment was calculated from the load cell data and was plotted against the averaged and full-wave rectified EMG data. A linear relationship was observed with an intercept on the extensor moment axis (M\textsubscript{e}) that indicated that flexion moment was resisted by passive tissues. The dependence of M\textsubscript{e} was further studied by having the subjects repeat the isometric pulls at varying amounts of trunk flexion. The isometric pulls showed that, on average, 3- to 4-fold increases in M\textsubscript{e} occurred for the full range of trunk flexion. The isometric pulls showed that, on average, 3- to 4-fold increases in M\textsubscript{e} occurred for the full range of trunk flexion.

Others too have contributed to the understanding of the biomechanical consequences of the FRP during upright trunk flexion. Gupta\textsuperscript{19} measured differences in myoelectric signals by experimentally inducing abnormal combinations of hip and trunk movements by having subjects bend forward with the pelvis held against a wall, which prevented posterior migration of the pelvis and limited the pelvic component of trunk flexion. In addition, subjects bent forward with weights tied posteriorly around the iliac crest, restricting pelvic movement, and with weights in their hands. During these tasks, EMG signals were obtained from surface electrodes placed on the erector spinae (at L3), rectus abdominis, hamstrings, and gluteus maximus muscles. A video camera captured the motion of the trunk flexion via markers placed on each subject and video was synchronized with the EMG data. The FRP was observed in all 25 subjects examined during quiet forward bending at 57% and 84% of mean maximum hip and vertebral flexion, respectively. Strong electrical activity was observed in the hamstrings muscles of all subjects during the flexion task, and 18 of 25 subjects showed mild activity of the gluteus maximus muscles during trunk flexion. The FRP was seen to come much earlier (P < .001) at an average hip and vertebral flexion of 28% and 75%, respectively, of the maximum flexion value when forward bending was performed with the buttocks held against the wall. This finding is consistent with the additional stability provided to the pelvis, thereby creating greater passive forces earlier in the flexion range of motion, further supporting the rationale for myoelectric silencing of the erector spinae muscles. The addition of weights placed either anteriorly or posteriorly was found to increase the tensile torque about the spine, requiring the erector spinae to remain active through a longer range of motion until the extension torque by the posterior vertebral elements is increased proportionally enough to reach equilibrium.

In another work, to understand the contributions of other muscles during trunk flexion, Andersson et al.\textsuperscript{19} used fine wire electrodes to measure the electrical activity of the quadratus lumborum (QL), the deep lateral and superficial medical lumbar erector spinae, the psoas, and the iliacus muscles during the FRP. The QL showed an increased involvement from erect standing to full forward flexion of the trunk. In the latter part of the trunk flexion, a cessation of activity was observed for the superficial medial portion of the lumbar erector spinae but not in the deep lateral lumbar erector spinae muscles. Based upon the results of this study, the common interpretation that extensor torque required at the lumbar spine to balance gravity being accomplished only by the passive tissues (ie, disks and ligaments) was found not to be entirely true. Muscles such as the QL and deep lateral erector spinae indeed assist in the load sharing for spinal stability in the flexed trunk posture. Knowledge of muscle activation patterns during trunk motions assists biomechanists and clinicians in understanding mechanisms of injury and providing direction in training and treatment regimens.

Clinical Studies on the FRP

Bridging the gap between the biomechanical research and clinical studies of the FRP was perhaps best accomplished by Kaigle et al.\textsuperscript{24} who simultaneously quantified the muscle activation patterns, the kinematic behavior of the lumbar motion segment, and the overall trunk flexion during dynamic flexion-extension in 7 patients with CLBP and 6 asymptomatic control subjects. Kaigle and coworkers\textsuperscript{24} performed invasive measurements in vivo by attaching interosseous pins into the spinous processes of lumbar motion segments. To the pins, a linkage transducer system allowed for dynamic measurement of intersegmental lumbar motion in the midsagittal plane with accuracies of 0.4° and 0.14 mm for rotatory and translatory movements, respectively.\textsuperscript{25} Electromyographic activity of the erector spinae muscles was simultaneously recorded from surface electrodes placed bilaterally at the level of L3 through L4.

With this experimental setup, subjects were asked to stand erect in quiet stance followed by bending forward in the sagittal plane to maximum trunk flexion, followed by extension to the upright neutral posture. Significant differences between the patients and asymptomatic subjects were observed throughout the flexion tasks. In the asymptomatic group, the authors found a 78% decrease in the root-mean-square EMG activity at full flexion, indicating the FRP. Conversely, in the patient group, 4 of the 7 patients showed no decline in erector spinae muscular activity at full flexion, and in the remaining 3 patients, two showed a less than 18% reduction in activity, whereas the third patient experienced a 60% reduction in activity at full flexion. Overall, the mean reduction for the patient group was 13%. Kaigle et al.\textsuperscript{24} further reported that the patients with CLBP
were significantly limited in their ability to flex and extend the trunk compared with the asymptomatic group.

Data obtained from the spinal linkage system revealed that sagittal rotation and axial shear translation were significantly less in the patients than in the control subjects. The authors reported that the restricted intervertebral motion in the patient group as well as trunk flexion may have been due to the persistent activation of the musculature. In this manner, the muscles would behave more as stabilizers to compensate for joint laxity in an injured or diseased motion segment to protect these spinal structures from movements that may cause pain and/or further tissue damage. Because subjects in both groups were able to flex greater than 70° previous study has shown that the FRP is seen approximately at this trunk angle in asymptomatic subjects, it is likely that the lack of the FRP in the patient group was not due to insufficient trunk flexion.

It is likely that the muscular alterations responsible for the absence of FRP in persons with back pain is related to their clinical status, namely, pain and dysfunction. The underlying rationale lies in the sensory innervation of spinal ligaments and their direct interactions with the surrounding spinal musculature. Although ligament has been traditionally categorized as a mechanical structure responsible for joint stability, they have equally important sensory functions. Neuroanatomic and neurophysiologic studies show that spinal ligaments and intervertebral disks are endowed with mechanoreceptors and nociceptive afferents to signal joint loads, motion, and the presence of inflammation in the latter. Ligamentomuscular reflexes have subsequently been established between the spinal ligaments and disks, and surrounding muscles, which act to directly or indirectly modify the load imposed on spinal ligaments through muscles. Such reflexes may be inhibitory or excitatory as may be required to preserve joint stability: inhibiting muscles that destabilize the joint or increasing antagonist coactivation to stabilize the joint. In fact, Solomonow et al recently described a neuromuscular disorder composed of 5 distinct components associated with static loads imposed to the spine. This model sheds light on theoretical mechanisms supporting the lack of FRP in patients with low back pain. The first component consists of a gradual decrease in reflexive muscular activity directly related to the creep developed in the ligaments, eliciting a shift in the sensory trigger threshold of the reflex. The second component consists of spams observed during the static loading period, caused by microdamage in the ligamentous collagen fibers and subsequently relayed nociceptively. The third component was observed in the first hour of rest after a static loading period, expressed as a transient hyperexcitability of reflexive muscular activity. Fourth, a relatively prolonged reflex muscular hyperexcitability (2- to 3-fold) that gradually increased from the second to sixth hour of rest after static loading has been observed correlated to the consequent ligamentous inflammation. The fifth component is the slow exponential recovery of the EMG to its normal (initial) level as rest time progresses.

The neuromuscular model as described by Solomonow et al may indeed represent the musculoskeletal dysfunction experienced by those with back pain and absent FRP. More recently, the effects of static flexion-relaxation on paraspinal reflex behavior have been reported. Reflexes showed a trend toward increased gain after a period of flexion-relaxation and were increased with trunk extension exertion. Other studies have investigated the reliability of detecting the FRP and its ability to discriminate between subjects with back pain and asymptomatic control subjects. Watson et al assessed the test-retest reliability of the FRP measure in a group of patients with CLBP (n = 11) and further compared the results between a group of healthy control subjects (n = 20) and a group of patients with CLBP (n = 70). In their study, Watson and coworkers developed a flexion-relaxation ratio of the activity during the neutral and fully flexed positions by taking the mean rectified mean square (RMS) EMG values during a 15-second resting standing period, RMS of the maximal activity during 1 second of activity during forward flexion, RMS activity for 1 second in the fully flexed position, and the RMS for the maximum activity during the reextension movement. Using a repeated measures design of 2 testing sessions 4 weeks apart, intersession reliability was found to range between 0.81 and 0.98 for EMG activity assessments. In the second experiment, the flexion-relaxation ratio was compared among the patients with CLBP and control subjects. The researchers reported that the flexion-relaxation ratio was significantly greater in the fully flexed trunk position in the CLBP group than in the control subjects, clearly discriminating the patients from the healthy control subjects in this study.

In related work, Shirado and coworkers also found that the FRP could discriminate between patients with chronic back pain and healthy subjects in their study of 25 healthy subjects and 20 patients with CLBP. Full wave rectified and averaged values of the EMG signal amplitudes were derived for the neutral upright and trunk-flexed postures, and based upon the averaged values of EMG activity during the relaxed erect position, changes (percentage) of EMG activities were calculated and compared between the groups. The authors found that all 25 healthy control subjects exhibited the FRP. Conversely, no patients with CLBP revealed the FRP during the flexed posture. Because of the significant difference in erector spinae EMG activity observed between the groups and a time lag between trunk and hip motion being greater in patients than in control subjects, Shirado et al concluded that neuromuscular coordination between trunk and hip function is altered among patients with CLBP.

Another study examining 40 patients with chronic back pain and 40 control subjects also showed a lack of FRP in patients with low back pain. In this work, Ahern et al noted that pain behavior, more specifically, guarded movements, was significantly correlated to the FRP. Adding to this
line of investigation of assessing the FRP, Mannion et al\textsuperscript{5} investigated 148 patients with CLBP and found that 55% showed no relaxation of the erector spinae muscles at L5 in the fully flexed position. Admittedly, the variance between the results of Mannion et al\textsuperscript{5} and those of studies previously cited indeed may lie in their choice of electrode location and methodology of EMG analysis. Nevertheless, this study corroborates the findings of others revealing an absence of the FRP among patients with CLBP.

Recently, Neblett and coinvestigators\textsuperscript{15} performed a two-part investigation of 12 asymptomatic subjects in an intraindividual repeated-measures protocol to examine the reliability of EMG readings of the FRP and further compared 54 patients with CLBP for differences in FRP responses before and after a spinal rehabilitation program. The authors reported that the ability of clinicians to measure range of motion and EMG reliably during the flexion task was high ($r > 0.92$, $P < .001$). Of further interest, all asymptomatic subjects achieved the FRP at mean EMG signals of 2.3 $\mu$V. Only 30% of the patients with CLBP achieved the FRP pretreatment; however, after a 7-week rehabilitation program, 94% of the patients with CLBP were able to achieve the FRP. This study not only showed the reliability of the FRP assessment procedure but also was the first study to systematically show that an absence of the FRP in patients with CLBP could be corrected with treatment.

**Conclusions**

The biomechanical consequence of the FRP is to accommodate the transfer of loads to the passive elements of the spine to achieve equilibrium. From a review of the biomechanical literature relevant to the FRP, the presence of the FRP during trunk flexion represents myoelectric silence consistent with increased load sharing of the posterior discoligamentous passive structures.\textsuperscript{7} The myoelectric silence of the erector spinae muscles per se, however, may not mean that these muscles are not providing forces themselves from passive stretching as previously thought\textsuperscript{6} nor that other lumbar spinal muscles do not assist in load sharing, namely, the QL and deep erector spinae muscles.\textsuperscript{19} The myoelectric silencing of the erector spinae muscles in the trunk flexion posture, although not exclusive, may be indicative of increased load sharing on passive structures,\textsuperscript{6,7,18,23} which tissues have been found to fail under excessive loading conditions,\textsuperscript{42} and showed to be a source of low back pain.\textsuperscript{43}

From the research reviewed, there is clinical significance to the presence or absence of the FRP. A number of studies have shown differences in the FRP between patients with CLBP and healthy individuals, and the reliability of the assessment. Persistent activation of the lumbar erector spinae musculature among patients with back pain represents the body’s attempt to stabilize injured or diseased spinal structures thereby protecting them from further injury and avoiding pain. Noteworthy is that few clinical tests show 100% correlation to the disease or disorder. The studies that show differences in the presence of the FRP among patients and control subjects are encouraging for this type of clinical assessment and suggest that assessment of the FRP is a valuable clinical tool to aid in the diagnosis and treatment of patients with CLBP. Further study into the response of different patient populations, including those patients with acute low back pain and sciatica, and the effects of different rehabilitation strategies will serve to improve the body of knowledge relevant to the clinical utility of the FRP.

**References**


LABRAL INJURIES OF THE HIP: A REVIEW OF DIAGNOSIS AND MANAGEMENT

Matt Schmerl, MChiro, a Henry Pollard, Grad DC, MSportSc, PhD, b and Wayne Hoskins, MChiro c

ABSTRACT

Objective: To report the current knowledge of the diagnosis and treatment of acetabular labral tears.

Methods: A search of the MEDLINE, CINAHL, and Science Direct indexing systems (1966 to September 2004) was conducted using the following key indexing terms: labrum, labral, hip, acetabulum, injury, and treatment. One hundred eighty-six publications were sourced using this methodology and were considered in this review. The literature was sorted according to publication date and relevance.

Results: There is a small amount of literature on the topic of labral lesions. This is particularly true of the use of conservative (manual therapy) methods for the treatment of labral lesions. The literature on surgical diagnosis and management is more mature; however, longer-term follow-up studies are required to conclusively show the benefit of surgical intervention.

Conclusions: Early diagnosis is important as labral tears may be linked to the progression of hip osteoarthritis. Initial treatment consisting of partial weight-bearing may respond if initiated early. Arthroscopy currently represents the gold standard in both the diagnosis and treatment of labral tears. Future research must investigate the long-term outcomes of partial labrectomy, as well as the efficacy of conservative approaches to care. (J Manipulative Physiol Ther 2005;28:632.e1 -632.e8).

Key Indexing Terms: Hip; Hip Injuries; Athletic Injuries; Labral Tear; Chiropractic

Only recently has the acetabular labrum been recognized within the orthopedic literature as a source of hip pain.1 Acetabular labral tears are the most common cause of mechanical hip symptoms.1 The multitude of diagnostic possibilities associated with anterior hip pain highlights the importance of an in-depth knowledge on behalf of the clinician to make an accurate diagnosis (Table 1).

The first report of a tear to the acetabular labrum was made in 1957 after Peterson2 described two cases of labral tears associated with irreducible posterior hip dislocation. The experience of Damerons3 in 1959 reported similar findings. The first report of nontraumatic tearing of the acetabular labrum was by Altenberg4 in 1977, as he described two cases of labral tears successfully treated by resection of the torn fragments. Suzuki et al5 described the acetabular labrum tear arthroscopically for the first time in 1986.

Practitioners are often faced with patients presenting with anterior and/or posterior hip pain on weight-bearing twisting movements of the leg. Often, the practitioner will consider a lumbosacral, thoracolumbar, or sacroiliac origin for these complaints. In older patients, an osteoarthritic origin may be expected. However, labral injury of the hip may present in the younger person who does not appear to be suffering spinal sources of pain or from degenerative hip causes of pain. The aim of this article is to report the current knowledge in the diagnosis and treatment of labral tears, including history, physical examination, and imaging, and to highlight clearly the criteria for conservative and surgical management.

METHODS

A search of the literature (1966 to September 2004) was conducted using the following databases: Medline, CINAHL, and Science Direct. The following key indexing terms were used: labrum/labral AND hip provided 186 responses, and when further words (acetabulum, injury, treatment) were added to the search criteria, the number of

632.e1
The labrum is a fibrocartilagenous structure composed of radially orientated collagen fibers attached to the osseous rim of the acetabulum. It is continuous inferorally with the transverse acetabular ligament. The labrum has been found to be larger in children than in adults, and the degree of coverage by the acetabulum and the labrum together is slightly larger in adults. The labrum is most often triangular in cross section. Aydinjoz and Ozturk performed comparative magnetic resonance imaging (MRI) of the labrum of both hips in 180 asymptomatic volunteers to find that 69% had a labrum with a triangular cross sectional area. Other studies show variable low percentages of round- and flat-shaped labrums in the anterior, superior, and posterior aspects. The labrum is thinner anteriorly and thicker posteriorly. There have been various reports of “absent” labrums after MRI. Incidence varies between 2.5%, 10%, and 14%. The significance of these findings is unknown.

In the study of 55 embalmed and 12 fresh frozen hips (mean age, 78 years) by Seldes et al, the acetabular labrum merged with the articular cartilage of the joint surface through a transition zone of 1 to 2 mm. A consistent thin tongue of bone extended from the edge of the bony acetabulum into the substance of the labrum. The labrum was found to attach firmly to the articular side of this bony extension via a zone of calcified cartilage. The labrum attached to the articular side of this bony extension via a zone of calcified cartilage with a well-defined tidemark. The outer surface attachment to the bony extension did not have a zone of calcified cartilage. They also noted a narrow synovial lined recess separating the labrum from the capsule, deepest superiorly and shallow inferiorly.

Although Seldes et al report a group of 3 to 4 blood vessels in the substance of the labrum, this observation has not been replicated. McCarthy et al performed immunohistochemical staining to confirm the blood supply of the region. They confirmed there was no evidence of penetration of vessels from the acetabular bone into the labral substance. In addition, they showed there were no gaps in vascular supply and no regions of relative hypovascularity. It was confirmed that the vascular supply of the acetabulum came from the obturator artery and superior and inferior gluteal arteries. The synovial tissue of the labrum capsular sulcus was highly vascular as was the outer surface of the acetabulum. Vessels penetrated as far as the junction between the bony acetabulum and the labrum.

In a study of 24 human acetabulae (mean, 64.8 years), the presence of nerve endings in the acetabular labrum was confirmed. It was further postulated that these nerve endings are indeed involved in nociceptive and proprioceptive mechanisms.

The exact function of the acetabular labrum is not fully understood. A study by Konrath et al concluded that removal of the labrum does not significantly increase the pressure between the acetabulum and femoral head and may not predispose to osteoarthritis. The study was well executed; however, the methodology, involving the use of pressure sensitive films and static loading conditions, may not be most reflective of the true clinical situation. A study

**Table 1. Differential diagnosis of labral injury causing hip pain**

<table>
<thead>
<tr>
<th>Category</th>
<th>Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contusion (especially over bony prominences)</td>
<td></td>
</tr>
<tr>
<td>Strains</td>
<td></td>
</tr>
<tr>
<td>Athletic pubalgia</td>
<td></td>
</tr>
<tr>
<td>Osteitis pubis</td>
<td></td>
</tr>
<tr>
<td>Inflammatory arthrites</td>
<td></td>
</tr>
<tr>
<td>Piniformis syndrome</td>
<td></td>
</tr>
<tr>
<td>Snapping hip syndrome</td>
<td></td>
</tr>
<tr>
<td>Bursitis (trochanteric, ischiogluteal, iliopsoas)</td>
<td></td>
</tr>
<tr>
<td>Osteoarthritis of femoral head</td>
<td></td>
</tr>
<tr>
<td>Avascular necrosis of femoral head</td>
<td></td>
</tr>
<tr>
<td>Septic arthritis</td>
<td></td>
</tr>
<tr>
<td>Fracture or dislocation</td>
<td></td>
</tr>
<tr>
<td>Tumors</td>
<td>Benign (simple bone cyst, osteoid osteoma, osteochondroma, fibrous dysplasia)</td>
</tr>
<tr>
<td></td>
<td>Malignant (Ewing’s sarcoma, osteogenic sarcoma)</td>
</tr>
<tr>
<td></td>
<td>Hernia (inguinal or femoral)</td>
</tr>
<tr>
<td></td>
<td>Slipped femoral capital epiphysis</td>
</tr>
<tr>
<td></td>
<td>Legg-Calve-Perthes disease</td>
</tr>
<tr>
<td></td>
<td>Referred pain from lumbosacral structures and the sacroiliac joint</td>
</tr>
</tbody>
</table>
by Tan et al, looked at 55 embalmed hips and calculated the acetabular articular surface area and volume with and without the labrum. They showed that structurally, the labrum provided an additional 28% surface area for the femoral head articulation. The study was unable to confirm how much, if any of the 28% is actually weight-bearing. It was also found that the labrum deepened the socket by an average of 5 mm. It has been suggested that the labrum may become a weight-bearing structure in the extremes of hip range of motion (ROM). It is at these extremes of range where labral tearing might occur. Another possible role of the labrum is to seal the joint, providing stability by allowing atmospheric pressure to aid in keeping the joint reduced.

Labral Tears

Data from cadaveric studies indicate that labral lesions are a common occurrence. Seldes et al, in a study of 55 hips, found 96% (53 of 55) to have gross labral tears. McCarthy et al explored 54 acetabulae and found 93% to have at least one labral lesion. Byers et al, in their investigation of 365 hips, found 88% of patients older than 30 years to have labral detachment from the articular cartilage. This so-called bucket handle lesion in which the labrum separates from the acetabular rim at the labral-cartilage junction was also most prevalent in the study of McCarthy et al.

It is generally accepted that most labral tears occur in the anterior, anterior-superior, and superior regions of this acetabulum. Fitzgerald noted a 92% incidence of anterior or anterosuperior location of tears in 55 active adult patients reporting a slipping or twisting injury with catching-type pain. Seldes et al found 74% of tears to be in the anterosuperior quadrant. Byrd also reported a majority of lesions to occur at the anterosuperior portion. Interestingly, posterior tears of the labrum appear to have a higher incidence in Asian populations. McCarthy et al report that these posterior tears in Asian populations are associated with hyperflexion or squatting motions. Ikeda et al, in their arthroscopic investigation of 7 young patients reporting pain from frank trauma, found all but one labral tear to be posterosuperior in location. Further studies with larger sample sizes are needed to decipher whether Asian populations do, in fact, have a higher incidence of posterior labral tears. Patients with known posterior subluxation or dislocation also most frequently have posterior tears. The most common of all lesions to occur in the anterior labrum is what McCarthy et al term the watershed lesion. This is the typical appearance seen after minor trauma in which there is an anterior labral tear present concurrently with anterior acetabular chondral injury.

There are several reported ways to classify labral tears. In this discussion, labral tears are classified according to location, etiology, and morphology. In respect to location, tears are classified as anterior, posterior, or superior (lateral). In respect to etiology, McCarthy et subclassify tears as degenerative, dysplastic, traumatic, and idiopathic. Degenerative tears may involve one or more regions of the labrum and can be seen in inflammatory arthritides. Labral tears and fraying are almost universal in patients older than 60 years. It has been suggested that labral tears predispose to osteoarthritis of the hip joint. If, in fact, the labrum is a weight-bearing structure, one would suspect that labral disruption would lead to nonsymmetric force distribution on the acetabulum and femoral head. It would be intuitive that such changes would lead to cartilage erosion, loss of joint space, sclerosis, and degenerative joint disease. This is an example of a labral tear causing degenerative joint disease; however, the high incidence of labral tears in patients older than 60 years may suggest that tearing is part of the normal aging process. Further studies are needed to discern whether all labral tears are symptomatic and whether tears are a part of the tissue aging process.

There have been many studies linking the presence of dysplasia to tears of the acetabular labrum. Dysplasias such as Legg-Calve-Perthes disease and congenital hip dysplasia (CHD) are now well known to predispose to labral tearing. McCarthy et al found tears to occur most frequently anterior in this population; however, they can be posterior or diffuse. They also found the labrum to be hypertrophic in the anterior portion causing infringement upon the anterior acetabulum. They conclude that the hypertrophy and tearing most likely cause impingement of the labrum between the acetabulum and femoral head, accounting for the mechanical symptoms frequently seen in this population. Patients in this category typically experience pain, clicking, and locking sensations of the hip joint.

Labral tears with a known traumatic onset are usually confined to a particular region of the labrum depending on the forces involved. Axial loading of the femoral head into the acetabulum and posterior dislocation lead to posterior labral tears. Anterior labral tearing normally results from minor trauma without dislocation. Hyperextension combined with femoral external rotation is the injury pattern most commonly associated with the acute presentation of acetabular labral tears. It is thought that the labrum takes on a weight-bearing role at the extreme of motion with excessive forces leading to tearing. Sports involving repetitive twisting motions and movements to end-range hyperflexion, hyperextension, and abduction are at greater risk.

Tears of an idiopathic etiology do not fall into the above categories and are found commonly in athletes with intractable hip pain and occupational related hip pain with no evidence of trauma.

Lage et al, in arthroscopic evaluation of 267 hips provided a morphologic classification of labral tears. Radial flap tears and radial fibrillated tears involve the free margins of the labrum and are the most commonly encountered. Longitudinal peripheral tears of various length are seen at the acetabulum-labrum junction, whereas unstable tears followed no real pattern but caused mechanical symptoms.
Diagnostic Considerations

It is known that dysplastic conditions of the hip are correlated with labral tears; therefore, it is important to ask about childhood hip problems such as Legg-Calve-Perthes disease, CHD, and slipped femoral capital epiphysis. Questions targeting the location, duration, frequency, and distribution of pain as well as aggravating and relieving factors are of critical importance. True intra-articular hip pain will cause pain in the groin area with occasional radiation to the knee. A patient with a labral tear will typically present with pain and mechanical symptoms. The pain is usually confined to the anterior groin but may also refer to the proximal greater trochanter or buttocks. There is frequently a painful clicking sensation or unpredictable “locking” of the joint. It is common for the symptom presentation to be considerably more subtle, characterized by dull, activity-induced, or positional pain that typically fails to improve with time. Patients with acute tears often recall some form of twisting injury, slip, or fall. In the study of 55 patients by Fitzgerald, only 30 experienced trauma to the lower extremity. Thus, in patients with an insidious onset of symptoms and signs, it is important to determine whether there is a history of repetitive rotational motion strain.

Sports that entail repetitive hyperextension and external rotation, such as soccer, skating, hockey, tennis, and golf, may predispose susceptible individuals to labral tears. Other conditions mimicking the clicking associated with labral tears are the various snapping hip syndromes around the hip due to either intra-articular or extra-articular conditions. Intra-articular causes include loose bodies, synovial chondromatosis, hip subluxation, and osteocartilagenous exostoses. Extra-articular causes include the posterior border of the iliotibial band over the greater trochanter, the iliopsoas tendon over the iliopectineal eminence, or the iliofemoral ligaments over the femoral head.

Differential Diagnosis

The differential diagnosis of hip pain of labral origin includes local musculoskeletal structures, dysplasias, tumors (benign and malignant), herniations, and referred pain from the abdomen, sacroiliac joint, thoracolumbar spine, and other structures. A list of differential diagnoses can be found in Table 1.

Physical Examination

The examination of a patient starts as the patient stands from their seat in the waiting room and walks to the examination room. Intra-articular disorders may be manifested by a Trendelenberg gait. If the hip is affected, the weight is lowered carefully on the affected side and the knee bends to absorb the shock. The step length on the affected side may be shorter than the normal side to minimize pain. Observation of the resting hip in flexion, abduction, or external rotation may indicate acute synovitis or joint effusion. Range of motion findings are often unrewarding. Fitzgerald has described a provocative maneuver to differentiate anterior from posterior labral tears. The sharp, catching pain with or without a click was reproduced in 54 (of 55) patients with the following: the hip is initially caught into acute flexion, external rotation, and full abduction, then extended with internal rotation and adduction. Pain with or without a click indicates an anterior tear. Moving from a fully flexed, internally rotated, and adducted position to an extended, abducted, and externally rotated position reproducing a sharp pain with or without a click indicates a posterior tear. McCarthy et al describes the McCarthy sign (with opposite hip flexed, the affected hip is extended first with internal then external rotation) in which a labral tear will reproduce a painful click. Suenaga et al in their study of 60 dysplastic hips, found the maximum flexion internal rotation test to be positive in 70% of patients with posterosuperior labral tears.

Imaging

Plain x-ray is effective at diagnosing many intra-articular disorders. Signs of Legg-Calve-Perthes disease, CHD, and slipped femoral capital epiphysis are all easily diagnosed on plain film. Plain film x-rays are not useful in detecting labral tears, chondral defects, or unmineralized loose bodies. Routine measurements, including the center-edge Angle of Wiberg, should be made to assess for acetabular dysplasia.

Conventional MRI is not thought to be of use in the visualization of labral tears. Diagnostic confidence is limited by the normal variability in labral size and shape, by the joint capsule collapsed against the acetabular rim, and by the difficulty in distinguishing tears of the labrum from pseudotears caused by the normal labral articular cartilage interface. In an MRI study of 180 asymptomatic hips, significant differences between subjects with respect to labral shape were reported. It has also been documented that there is an age-related intensity increase on T1-weighted and T2-weighted gradient echo magnetic resonance images of the acetabular labrum. These asymptomatic patients may or may not exhibit labral tears. In the 2003 review of Narvani, it is stated that MRI features that may suggest a labral tear include irregular labral shape, a nontriangular labrum, a thickened labrum with no labral recess, labrum with increased signal intensity on T1 images, and labral detachment from the acetabulum. From the previously mentioned studies these findings may all be normal variants. In addition, Byrd states that partial separation of the labrum from the lateral aspect of the acetabulum has been found to be a normal anomalous variation. In the study of 56 hips by Fitzgerald, only 3 were found to have a tear from MRI. Most patients did, however, show a low-grade effusion. Magnetic resonance imaging was effective in eliminating osteonecrosis, pigmented villonodu-
Treatment of Labral Tears

The diagnosis and treatment of acetabular labral tears are important not only for the relief of symptoms, but also to alter the progression of osteoarthritis. A trial of conservative treatment is indicated in cases of labral tears. To date, there has been no research on the efficacy of hip manipulation or mobilization in the treatment of labral injuries, despite indirect evidence that these procedures may be beneficial. In contrast, a small study has investigated the effect of hip manipulation or hip stretching on hip joint ROM in normal young adults or adult low back pain sufferers. Although these studies have shown improved hip joint ROM after treatment, the results cannot be extrapolated to a population of subjects suffering a labral tear. However, as manipulation has shown an improvement in function, further study should investigate the effect of manipulation on those with labral tearing. This is especially true for those that suffer from some catching or locking as a component of their symptom spectrum.

The labrum is an avascular structure and, therefore, receives nutrition and clears waste via imbibition. With restricted ROM comes lower rates of imbibition and compromised nutrition. Manual procedures aimed at restoring the rate of nutrient flow through increased joint ROM may promote labral healing. In cases of entrapped or folded labral portions, manual procedures, including hip joint tractional maneuvers, may be of benefit. Conservative treatment will usually consist of nonsteroidal anti-inflammatory drug administration and protected (partial) weight-bearing for 2 to 4 weeks. Ikeda et al reported 6 of 7 patients to respond to this regime. Fitzgerald, who had all patients (55 in total) partial weight-bearing for 4 weeks, found 7 patients experience complete relief of pain without recurrence. Hase and Ueo had two of their patients managed with bed rest and traction for a few weeks to find that both failed. It is unfortunate that most nonoperative interventions fail and patients report continued symptoms of painful catching particularly with activities. Conservative treatments often fail because labral tears are actually diagnosed on average of 2 to 3 years after injury. This delay in diagnosis is 2-fold: many patients cannot recall an antecedent injury or event, and many patients are initially misdiagnosed. It would be useful for further studies to investigate the effectiveness of protected weight-bearing on labral tears of recent onset. Other conservative approaches also warrant further investigation.

Rehabilitation

There are currently no studies investigating the effects of rehabilitation on pre- or postarthroscopic patients. However, graded rehabilitation is generally prescribed to arthroscopy patients for several reasons. The authors here recommend a conservative trial of care presurgically. The search for biomechanical predisposing factors may be of benefit in cases of labral tears. Janda has described a predictable pattern of muscular imbalance in the pelvis, known as the lower crossed syndrome. Tightness of the hip flexors and lumbar erector spinae and weak, inhibited gluteal and abdominal muscles characterize lower crossed syndrome. The resultant imbalance leads to anterior pelvic tilt, increased hip flexion, and a hyperlordosis of the lumbar spine. Hip flexion contracture might lead to increased weight-bearing upon the anterior acetabulum and labrum predisposing to tearing. Many patients presenting in the primary setting with low back pain due to lower crossed syndrome can be rehabilitated to correct these aberrant

Magnetic resonance arthrography (MRA) is a modality available to image labral tears, and the use of contrast medium within the hip capsule makes this an invasive procedure. The principle of the procedure relies upon capsular distension, thereby outlining the labrum with contrast (usually gadolinium) and filling any tears that may be present. Labral tears manifest as an abnormal linear extension of high-intensity gadolinium solution into the labrum, labral blunting, or detachment from the underlying bone. The addition of intra-articular contrast has been shown to greatly increase the specificity and sensitivity of diagnosing labral tears. The normal recesses occurring at the labral-cartilage junction can make the normal labral sulcus difficult to distinguish from a labral tear. In a study of 22 hips by Czerny et al, labral tears on MRA were confirmed surgically. In this study, MRA had a sensitivity of 90% and accuracy of 91%, compared with a sensitivity of 30% and accuracy of 36% with conventional MRI. Fitzgerald found a tear in 44 of 60 hips imaged with MRA. The injection of local anesthetic at the same time as arthrography was used to confirm the presence of intra-articular pathology or extra-articular source of pain. The average subjective pain relief (analog scale) experienced with marcaine injection at the end of the hip arthrography was 81% (range, 0%-100%). Although numerous studies have documented success of MRA in detecting labral tears, Hase and Ueo reported a somewhat different experience. Ten patients were considered normal at arthroscopy. They conceded they could not visualize the posterior portion of the labrum as the femoral head obstructed its view. Considering most tears were posterior in location, this result was not surprising. McCarthy et al stated that the specificity of detecting anterior tears was 83% specificity and 78% accuracy. The detection of posterior tears (20%) and lateral tears (11%) is poor.

Arthroscopy, as a diagnostic tool, is reserved for those patients with an intra-articular source of pain without a clear diagnosis after careful history, physical examination, and radiographic studies. Arthroscopy will be discussed further in the treatment of labral tears.
pelvic mechanics. If the patient is participating in high-risk sports, they may be in a position to distribute forces more evenly around the labrum and acetabulum, reducing the risk of tearing. Dirocco et al. have noted many impairments after arthroscopy; these include inflammation, pain, swelling, decreased joint mobility, altered muscle extensibility, impaired muscle strength, altered proprioception, and decreased muscle endurance. Rehabilitation protocols should be designed to address these impairments. Manual therapy should be included, given its effects of increasing joint mobility, stretching muscles, and increasing proprioception. It is important to maintain a flexible approach, however, as all patients will have varied needs. The initial period after surgery is focused to decrease pain, decrease effusion, prevent muscle inhibition, promote tissue nutrition and wound healing, maintain proper static joint alignment, increase awareness of joint protection, allow an independent and safe gait with assistive device on all level surfaces and elevations, and increase sitting tolerance.

An aquatic program is often beneficial allowing early joint mobilization and gentle strengthening in a reduced-weight environment. After initial pain relief, all efforts should be focused toward restoring hip ROM through muscle stretch procedures and capsular mobilization. Small accessory oscillation movements stimulate joint mechanoreceptors and assist in pain modulation. Faulty movement patterns should also be addressed while improving muscle strength and endurance. Griffin states that proprioceptive deficits routinely occur in conjunction with injury to the joint surfaces. This inhibits normal motor response and decreases neuromuscular stabilization of the joint. Thus, proprioceptive retraining is important to restore these deficits and assist in re-establishing neuromotor control. Altered gait and weight-bearing mechanics are likely to predispose to sacroiliac and lumbar spine dysfunction. The lumbopelvic rhythm must be assessed in all patients with altered hip mechanics. Spinal manipulative therapy is likely to be of use in those patients with sacroiliac joint and lumbar spine dysfunction. Abdominal core (transverse abdominus) stabilization exercises are also likely to be of use in restoring correct lumbopelvic biomechanics. With functional progression of the patient comes the introduction of activity-specific exercises and plyometric training with slow return to work or sporting activity. Patients with resection of a labral tear without evidence of articular cartilage involvement may progress more quickly to closed kinematic-chain, weight-bearing exercise and return more quickly to sports or to physically demanding jobs. Return to sport is usually possible in 2 to 4 months.

**Arthroscopy**

Arthroscopy currently represents the gold standard in both the diagnosis and treatment of labral tears. It is now considered minimally invasive owing to more advanced knowledge of the labrum as well as technique and instrument development. Visual inspection of all quadrants of the joint is possible with arthroscopy. Table 2 lists arthroscopic indications and contraindications.

For the procedure, the patient is usually placed in either the supine or lateral position, depending on which position the surgeon finds more comfortable. Huffman and Saffran and Byrd describe their preference of the supine position. McCarthy prefers the lateral decubitus position. Once the patient is positioned correctly, the operative leg is distracted in a longitudinal, slight lateral, and caudal vector in line with the femoral neck. Byrd describes 3 portals in the supine position. The anterior portal lies (on average) 6.3 cm distal to the anterior superior iliac spine and penetrates the muscle belly of the sartorius and rectus femoris before entering through the anterior capsule. The anterolateral portal penetrates the gluteus medius before entering the lateral aspect of the capsule at its anterior margin. The posterolateral portal penetrates both the gluteus medius and minimus before entering the lateral capsule at its posterior margin. The anterior capsule is best visualized through the anterior portal, whereas the posterior labrum is best seen.

### Table 2. Indications/contraindications of hip arthroscopy for a labral lesion

<table>
<thead>
<tr>
<th>Indications for hip arthroscopy</th>
<th>Contraindications</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Loose bodies</td>
<td>• Femoral neck stress fracture, pelvic insufficiency fracture</td>
</tr>
<tr>
<td>• Synovial chondromatosis</td>
<td>• Transient osteoporosis</td>
</tr>
<tr>
<td>• Labral tears</td>
<td>• Osteonecrosis and synovitis without mechanical symptoms</td>
</tr>
<tr>
<td>• Chondral flap lesions of the acetabular or femoral head</td>
<td>• Acute skin lesions</td>
</tr>
<tr>
<td>• Osteonecrosis of the femoral head</td>
<td>• Sepsis with osteomyelitis or abscess formation (open surgery)</td>
</tr>
<tr>
<td>• Ruptured or impinging ligamentum teres</td>
<td>• Joint ankylosis, protrusio limiting hip distraction</td>
</tr>
<tr>
<td>• Collagen disease (RA, SLE, etc) with impinging synovitis</td>
<td>• Advanced osteoarthritis</td>
</tr>
<tr>
<td>• Foreign body removal</td>
<td>• Obesity (relative contraindication)</td>
</tr>
<tr>
<td>• Crystalline hip arthropathy (gout, pseudogout etc)</td>
<td></td>
</tr>
<tr>
<td>• Capsular shrinkage (Ehler-Danlos, etc)</td>
<td></td>
</tr>
<tr>
<td>• After trauma (dislocation, Pipkin fracture, bullet etc)</td>
<td></td>
</tr>
<tr>
<td>• After total hip arthroplasty</td>
<td></td>
</tr>
<tr>
<td>• Intractable pain with positive physical findings</td>
<td></td>
</tr>
<tr>
<td>• Osteoarthritis</td>
<td></td>
</tr>
</tbody>
</table>
through the posterolateral portal. The initial portal placement is usually the anterolateral portal, as this is located in the so-called safe zone of the hip. In this position, there is limited chance of neurovascular compromise. Once in place, the anterolateral portal can assist in directing the remaining portals. Once labral tears are identified, the surgical portion of the operation is used, which involves minimal debridement of the torn labral segment. It should be noted that there is considerable confusion in the orthopedic literature regarding nomenclature of portals. It would be useful to report a standardized definition of each portal placement.

The short-term results of arthroscopic debridement of labral tears have been favorable. O’leary et al\textsuperscript{36} reported a 91% success rate in relief of symptoms in a series of 86 patients undergoing arthroscopy. Santori and Villar\textsuperscript{37} reported more than two thirds of their patients having positive symptomatic relief. Byrd and Jones\textsuperscript{38} documented the Harris hip score as an outcome measure after arthroscopic labrectomy. It should be noted that the Harris hip score is traditionally used and was designed to assess the short-term outcome. Future research needs to focus on the long-term consequences of partial labrectomy. Because a section of the labrum is excised, the mechanical properties of the structure are altered. Although symptomatic relief is important, partial labrectomy could itself still lead to progression of hip osteoarthritis. Arthroscopic surgery still carries risks; various complications are noted in Table 3. Complications after hip arthroscopy are stated to occur in 1.6% to 5% of cases.\textsuperscript{35}

### CONCLUSION

Tears of the acetabular labrum are increasingly recognized as a source of hip pain, particularly in athletes. The clinical presentation may be similar to referred pain from lumbosacral or sacroiliac structures, making awareness of labral injuries valuable when considering a list of differential diagnosis. In the North American population, tears are most frequently anterior and are often associated with sudden twisting or pivoting motions. In the Asian population, tears are more frequently posterior and associated with hyper flexion or squatting motions. Because the data are very ethnic-specific, further studies may wish to use a wider ethnic mix. No radiographic study, including MRA, is able to diagnose labral tears with significant specificity or sensitivity. The clinical suspicion of a labral tear should result from thorough history taking as well as positive physical examination findings. Early diagnosis of labral tears is important as they may contribute to the progression of hip osteoarthritis. Arthroscopy may be beneficial for the short-term relief of symptomatic labral tears; however, long-term consequences are unknown. The long-term results of arthroscopic labral debridement need to be confirmed in future randomized controlled trials. No research has been conducted into manipulation or mobilization for treatment of labral injuries. Given the indirect evidence for benefit, this should be investigated in clinical trials.

### REFERENCES

CASE REPORT

PATHOLOGICAL CERVICAL FRACTURE AFTER SPINAL MANIPULATION IN A PREGNANT PATIENT

Alfred Schmitz, MD,a Goetz Lutterbey, MD,b Lars von Engelhardt, MD,a Marcus von Falkenhausen, MD,b and Michael Stoffel, MDc

ABSTRACT

Objective: To present the rare case of a displaced odontoid fracture after manipulative treatment.
Clinical Features: A 37-year-old, 15-week pregnant patient was referred with acute neck pain and a diffuse paravertebral swelling that started after cervical manipulation performed by her general medical practitioner 5 days before. Because of pregnancy, a cervical spine radiographic series was not obtained before treatment. Magnetic resonance imaging revealed a displaced odontoid fracture associated with a pathological process in the vertebral body of C2 and a paravertebral hematoma on the left side from C2 to C4.
Intervention and Outcome: After initial halo vest immobilization, an anterior-posterior fusion of C1-C2 was performed. The histological analysis showed features of an aneurysmal bone cyst. The patient was discharged and had an undisturbed pregnancy and was without any neurological complications.
Conclusions: Because of the weakening lesion in C2, the spinal manipulation most likely caused the displaced odontoid fracture. Special imaging should be performed, preferably with magnetic resonance imaging, when a patient experiences significant new symptoms after cervical manipulation. (J Manipulative Physiol Ther 2005;28:633-636)

Key Indexing Terms: Manipulation, Spinal; Pregnancy; Fractures, Spontaneous

Spinal manipulation is a widely used and successful treatment option for reversible functional disorders of the cervical spine.1 Severe complications after manipulative treatment rarely occur.2,3 The complication rate for cervical spine manipulation is estimated to be between 5 and 10 per million manipulations.4 However, rare reported complications include vertebral artery dissection, cord or root injury, phrenic nerve injury, epidural hematoma, cervical disk rupture, and vertebral fracture.2,5-11
Vertebral fractures after spinal manipulation have been described in patients with osteoporosis and tumor disease, which are contraindications for manipulative treatment.5,7 But, as in the reported case, ruling out these contraindications is sometimes difficult without radiographic evaluation. A fracture of the cervical spine in association with manipulation and an underlying spinal tumor has not been reported previously.
Most contraindications to spinal manipulation are evident during a careful history and physical examination. Therefore, radiographic assessment is not routinely required for the diagnosis of spinal conditions treatable with manual therapy.12 In the presence of risk signs suggesting a possible underlying pathological condition, however, x-rays of the cervical spine should be performed. Still, like in the following case, contraindications for radiographic imaging, such as pregnancy, sometimes forbid radiographic assessment before commencing treatment.
We report the case of a pregnant patient with acute onset of neck pain after cervical manipulation. A magnetic resonance image (MRI) taken 5 days later revealed a C2 odontoid fracture that was displaced.

CASE REPORT

A 37-year-old, 15-week pregnant woman was admitted to our department with an acute onset of neck pain and a
progressive paravertebral swelling. She received a single, manipulative treatment of the cervical spine 5 days earlier. At that time, she received paravertebral injections by her general medical practitioner, who also did the manipulation. This treatment was initiated because of diffuse neck pain for the previous 2 weeks. She had no history of trauma or tumor and had an uneventful pregnancy to date.

During the clinical examination, a left paravertebral mass from C2 to C4 without signs of inflammation and a painful stiffness of the neck was identified. No neurological deficits were noted. Blood analysis showed a slightly increased C-reactive protein and a mild leukocytosis. Because of the acute symptoms and nature of the onset, a cervical MRI series was performed. This revealed a pathological type II odontoid fracture with ventral displacement of the odontoid process leading to compression of the spinal canal (Fig 1).

No signs of myelomalacia were observed. The paravertebral mass from C2 to C4 was a hematoma. Fig 2 shows the

**Fig 1.** MRI. The sagittal T2-weighted turbo-spin-echo (TSE) sequence with spectral fat saturation (SPIR) shows a displaced type II odontoid fracture and a hematoma in the prevertebral soft tissue (*).

**Fig 2.** MRI. The sagittal T2-TSE sequence after repositioning reveals a tumor in the vertebral body of the axis (*).

**Fig 3.** A sagittal reconstruction from a cervical spine CT shows clearly an osteolytic destruction of the vertebral body as the underlying cause of the fracture (*).
sagittal MRI sequence after repositioning, which revealed a tumor in the vertebral body of the axis.

Immediately after the MRI was reviewed, the patient was immobilized in a halo vest. For further surgical planning, a cervical computed tomography (CT) was carried out. The CT scan showed an osteolytic lesion of the odontoid base and the body of C2 (Fig 3). Subsequently, the pathological fracture was stabilized in a 2-step procedure including a dorsal fixation of C1/C2/C3 and an anterior resection of the pathological bone and spondylodesis with an autologous iliac crest implant. The histological analysis was performed and the tissue specimen was consistent with an aneurysmal bone cyst. There were no signs of infection or malignancy. The patient was discharged with her pregnancy undisturbed and without any neurological complications.

**DISCUSSION**

The patient had a bone-weakening lesion of C2 histologically determined to be an aneurysmal bone cyst. It is presumed that because of the underlying disease process, the cervical manipulation led to the pathological fracture of C2 or at least to the acute displacement of the odontoid causing the severe neck pain and the paravertebral hematoma.

Despite the increased pain and the paravertebral swelling after manipulative treatment, further evaluation with MRI was delayed for 5 days. The reason for the delay is presumably that the occurrence of the serious complication was not taken into consideration by the general practitioner. Because of the pregnancy, radiographic analysis was not carried out.

Spine fractures associated with manipulative treatment are very rare complications and are only reported with preexisting lesions such as osteoporosis, tumors, or metastases. Missed cervical spine fractures after trauma may also be a source for complications in the connection with spinal manipulation. A careful review of the radiographic examination is recommended in patients with previous cervical spine trauma to rule out contraindications for manipulation such as fracture or instability. Typical causes for missed contraindications after cervical trauma are initially incomplete or false-negative interpreted radiographs.

With a careful history and clinical examination, most preexisting lesions representing contraindications for manipulative treatment can be excluded. However, benign spinal tumors may not present with a typical history, and clinical examination might not reveal an osteolytic process in the cervical vertebra. Accordingly, only radiographic imaging will display such cervical lesion. On the other hand, cervical lesions due to benign spinal tumors, as reported in this case report, are extremely rare and routine radiographic imaging is not recommended in the premanipulative assessment. However, if the history or the examination reveals any risk signs, a radiographic evaluation should be performed before commencing cervical spine manipulation.

In the reported case, the histological study identified an aneurysmal bone cyst, which is a pseudotumoral hyperemic lesion of unknown etiology. Aneurysmal bone cysts are more prevalent than any other benign tumors in the spine with an incidence of 10% to 15%. In a report on 41 cases with aneurysmal bone cyst of the mobile spine, Boriani et al. found as constant clinical symptom a slow and gradual onset of pain. A palpable mass was present in 32% and neurological impairment in 10%. Of 14 (30%) cases involving the cervical spine, 3 had a pathological fracture. The diagnosis of an aneurysmal bone cyst can be achieved by radiography, MRI, and CT, but biopsy is mandatory for histological confirmation. Surgical resection, radiotherapy, and embolization are common treatment modalities for aneurysmal bone cysts involving the spine.

Because the patient was pregnant, an x-ray of the cervical spine was not performed. The general practitioner was not able to exclude the underlying contraindication for spinal manipulation. Because this is reported for the first time, it is presumed that a pathological odontoid fracture after spinal manipulation is an extremely rare complication.

**CONCLUSION**

A careful history and clinical examination may exclude most preexisting contraindications for manipulative treatment. If spinal manipulation is recommended for pregnant patients, physicians performing high-velocity, low-amplitude manipulation should consider the lack of radiographic imaging information in their clinical decision making. The current case report shows that special imaging should be performed, preferably with MRI, when the patient experiences significant new symptoms after cervical manipulation.

**REFERENCES**


THE BOURNEWORTH QUESTIONNAIRE: CAN IT BE USED TO MONITOR AND PREDICT TREATMENT OUTCOME IN CHIROPRACTIC PATIENTS WITH PERSISTENT LOW BACK PAIN?

To the Editor:

I am writing to dispute the conclusion that the Bourneworth Questionnaire (BQ) is not a useful instrument to monitor progress based on data that in the main, relate to properties of the instrument in predicting (and not monitoring) outcomes. This instrument was never designed to be used, nor as far as I am aware has ever been proffered, as either a diagnostic or a predictive instrument. Thus, to test it in this way, and then to make the leap and imply that it is also not useful as an outcome measure, seems to me to be unfair, biased, and unjustified.

Specifically, the BQ, when tested against the revised Oswestry questionnaire in this study, performed as well as the Oswestry in both responsiveness and predicting outcomes. A simple calculation of effect sizes from the data presented in table 3 illustrates comparable responsiveness and, by the authors’ own admission, no difference in predictive value (although by modifying the BQ, it was possible to predict “fairly well” all 4 of the outcomes tested).

The authors also investigated concurrent validity of the questionnaires. My interpretation was that the questionnaires agreed very well, with very small mean score differences at all time points. The authors stated that, “At all points in time, there was a small mean difference between the sum scores...” However, as they correctly observed, there was a wider scatter in the differences as the average increased (an observation not uncommon in these types of analyses). This later appeared as, “There is considerable disagreement between the... at all times meaning that one or both of the questionnaires produce unreliable data.” Finally, by assuming that the Oswestry is the criterion standard (a biased assumption because there is no way of knowing the “truth”), the authors imply that the BQ is unreliable. This is not supported by the data. Only well-designed replication studies can judge the reliability of an instrument. A final point of concern is that the authors failed to reference any rigorous validity studies of the translated versions of the questionnaires, thus calling into question the accuracy of the data collected by these instruments in either this or in any other studies.

I do agree with the authors, however, when they state that patients’ responses should be categorized using cutoff scores to enable calculation of numbers needed to treat. The authors may be interested to know that this has already been done for the BQ.3,4

The authors state that it is their opinion (and one I concur with) that the scientific and clinical world should not be further burdened with additional instruments unless they add to it. The BQ adds to it by virtue of being an outcome measure not only with robust psychometric properties comparable with other outcome measures but also one that is shorter and arguably more practical and one that can be used in the confidence that it measures a wide spectrum of outcomes. This view is reflected in the number of chiropractors, at least in the UK, who do find the BQ useful and use it routinely to monitor their patients’ progress. I trust that in spite of the contestable conclusions of this paper, they will continue to do so.

Jennifer E. Bolton PhD, MA Ed, ILTM, FCC(Hon)
Department of Research and Professional Development
Anglo-European College of Chiropractic
Parkwood Road
Bournemouth BH5 2DF, England

REFERENCES

RESPONSE TO LETTER TO THE EDITOR BY BOLTON

In Response:

There is a multitude of questionnaires for patients with low back pain (LBP). It has been recommended that two well-tested instruments be used in the future to allow comparison between studies and study populations.1

A new instrument, the 7-item Bournemouth Questionnaire (BQ), was developed, which appears promising in that it contains also a psychological dimension (lacking in the other two questionnaires) yet is shorter than the other two.4 However, the introduction of new questionnaires requires that these are better than the two already recommended.

Because the BQ might be a “better” questionnaire, we decided to take a closer look at it, specifically to find out if (1) all the 7 items in the BQ were relevant; (2) the responses to the BQ were similar to those of the Oswestry questionnaire; and (3) the BQ could predict treatment outcome.

We found that (1) anxiety and depression seemed to be largely irrelevant in our study population; (2) the same person could be categorized as very ill in one questionnaire and not
that ill in the other; (3) the BQ used, as recommended,\(^5\) as an index was not good at predicting outcomes.

We concluded that, “the BQ is not a useful instrument to identify baseline status, monitor progress, or predict the 1-year progress in chiropractic patients having persistent LBP. However, certain individual items are useful to predict specific outcomes.”

A letter to the editor brings forth the following points of disagreement:

1. We studied the BQ’s predictive and not monitoring properties. 
Response: If an instrument is not related to outcomes, why should it be used?
2. BQ does add to the main two recommended questionnaires because it has robust psychometric properties comparable with other outcome measures, that it is shorter and more practical, that it can be used in the confidence that it measures a wide spectrum of outcomes, and that it is already used by many chiropractors.
Response: In order to replace two already highly accepted outcome instruments this is not enough. You have to demonstrate that it is also better. Unfortunately, the results of our study did not support the latter.
3. The accuracy of our data is queried because there was no reference to any rigorous validity studies of the translated versions of the questionnaires into Norwegian.
Response: The BQ was translated into Norwegian and retranslated into English in a less elaborated manner than recommended by Beaton et al\(^6\). However, the questions in the BQ are straightforward and there are no obvious cultural differences between Norway and England, so the questions should be easily translated and understood.
4. The BQ was never intended to be a diagnostic or predictive instrument. To test it as such is “unfair”.
Response: Science is characterized by development. Therefore, one cannot prevent others to study the BQ from different angles. Whether this is “unfair” or not, surely, is irrelevant.

Our final conclusion is that until a questionnaire is proven to be better, it is better to follow the recommendations by Deyo et al\(^1\) to keep on using the Oswestry and Roland-Morris questionnaires to make the effect of LBP treatment comparable between patients and professions worldwide.

Kristian Larsen, PT, MPH
Musculoskeletal Research Unit
Orthopedic Department
Lægaardvej 12 7500, Holstebro, Denmark

Charlotte Leboeuf-Yde, DC, MPH, PhD
Research Professor
Back Research Center
Backcenter Funen and University of Southern Denmark
Denmark

References

False Negative Magnetic Resonance Imaging Results: A Report of 2 Cases

To the Editor:

Although case studies are typically designed to describe unusual findings (eg, a new intervention, a rare condition or a rare manifestation of a ubiquitous condition), Schneider has provided a well-written paper that describes a common occurrence in the context of an important clinical concept - no diagnostic test has 100% accuracy.\(^1\)

Diagnostic testing is an extremely important aspect of health care and accounts for approximately one-fourth of all ambulatory care expenditures in America.\(^2\) Diagnostic measures aim to identify individuals who could potentially benefit from intervention. But does every diagnostic test provide accurate answers? Does every test yield patient benefit?

Practitioners tend to think in terms of certainty, but healthcare issues are a matter of probability. From my experience in teaching evidence-based principles, practitioners interpret the results of tests, especially imaging, as either true positives or true negatives. Yet, all tests, including diagnostic imaging, have false positives and false negatives.

The false negative rate for MRI can be as high as 40% (40 patients out of 100 with a negative finding are really...
positive). The false positive rate for MRI can be as high as 57% (57 patients out of 100 with a positive finding are really negative). Moreover, sensitivity and specificity are not stable characteristics of a test, because they change with disease prevalence. As a rule, estimates of a test’s sensitivity are most accurate when disease prevalence is high, and estimates of specificity are most accurate when disease prevalence is low.

Although advances in diagnostic imaging have resulted in improved patient care for some diagnostic groups, these advances have also created risks of overestimating health problems, because small irrelevant “abnormalities” and anomalies can be detected. Thus, unnecessary interventions can be applied. Although the risks of the tests themselves may be relatively small, the cascade of subsequent events may quickly spiral out of control. Selecting only valid and reliable diagnostic tests could be an important strategy for controlling the clinical cascade. Understanding that a diagnostic test provides an estimate, not a guarantee that the patient was classified properly, can assist the practitioner in managing the patient and improving outcomes.

REFERENCES