



Angkor Children's Hospital, Cambodia. Orthopaedic overseas.

Desmond Brown, Boston, Massachusetts

The "Imagery" section of *Spine* is devoted to the artistic and imaginative qualities of our readers.

*Spine* invites you to submit your drawings, illustrations, or photographs, along with appropriate explanatory information, for consideration of publication within this section. Please forward two 5" × 7" copies of your work to: Dr. William A. Abdu, Coordinating Editor "Imagery," c/o *Spine*, Dartmouth College, 7254 Strassenburgh Hall, Hanover, NH 03755-3863. Tel: 603-650-1122 or 800-934-9379; Fax 603-650-1500.

## Harrington Lecture: The Future of Academic Spine Surgery: Challenges and Opportunities

David S. Bradford, MD

I wish to thank Dr. Jim Ogilvie and members and guest of the SRS for the privilege and honor of being your 2004 Harrington Guest Speaker. I first met Dr. Harrington in 1965 during my residency at NY Orthopedic, and through my subsequent association with Dr. Moe in Minneapolis, was greatly influenced by his revolutionary treatment of patients with spine deformities. His scientific interaction with Dr. Moe was indeed a mutually beneficial one. Dr. Harrington taught Dr. Moe his new instrumentation technique, and Dr. Moe correctly predicted that a fusion was still necessary, and taught Dr. Harrington the importance of a well-executed arthrodesis.

My subject for today is Academic Spine surgery. The term Academic Spine Surgeon might prove somewhat pejorative for are not we all by the nature of our membership in this society considered "academic?" So what is "academic?" From [www.word.net](http://www.word.net), it is defined as "an adjective implying hypothetical or theoretical and not expected to produce an immediate or practical result," "something marked by a narrow focus or display of learning especially its trivial aspects," or "an educator who works at a university." My comments will apply to the latter definition: a clinician, researcher, educator, and administrator based primarily, but not exclusively, at an academic medical center (AMC).

In my talk today, I'd like to begin by discussing what I feel are the primary issues facing spine surgery in general, the problems facing academic spine surgery and academic spine surgeons in particular, and finally recommendations and solutions. I will maintain that problems are only opportunities in disguise and indeed the opportunities are substantial. I will hope to convince you that the glass may be more than half full for those willing to take the plunge and willing to support the effort.

### ■ General Issues Facing Spine Surgery and Spine Surgeons

The issues and difficulties facing us as spine surgeons are well documented and are not dissimilar to those facing surgeons in other specialties: escalating malpractice premiums, managed care bureaucracy, decreasing reim-

bursement, increasing overhead cost, preauthorization requirements and approval for surgery, increased patient expectations, greater government oversight, HIPAA and SB1386 requirements, and the ever greater need for documentation that has taken us further away from what we do best . . . taking care of our patients.

At the same time, the explosive growth of spine surgery in North America over the past 5 years is a cause for reflection, if not concern. The national survey data indicate that the annual number of spinal fusion operations rose by 77% between 1996 and 2001.<sup>1</sup> From 2002 to 2003, cervical fusions increased by 12% and T/L fusions increased by 11%. In contrast, hip and knee replacements increased by 13% during the same period.<sup>2</sup> Spine fusion as we know is not an inexpensive operation. The average hospital bill is more than \$34,000, excluding professional fees. The more recent data for 2002 to 2003 presented by Orthopedic Network News report that combined procedures increased 34% and vertebral/kypoplasty at 20%. The use of bone morphogenetic protein (BMP) increased in this period to 14% of fusions; 50% of these procedures using BMP did not use a cage and 71% surprisingly used iliac bone graft with BMP! It is recognized that this information came from Medicare data composed of a very limited sample size.

We might as a research society take pause and attempt to understand why this rapid growth in fusion rate has occurred and more importantly whether our patients are being better served by this increase in number of fusions. This growth rate can no doubt be explained in part by technology advancements, an aging population, improved anesthesia, better diagnostic imaging, and wider and more variable indications for fusion surgery, particularly lumbar fusions as well as the increased number of spine surgeons.<sup>1</sup> Clearly, the regional differences in the prevalence of spine surgery in North America (United States) cannot be related to differences in disease patterns or failures in the conservative management of pain. Data would suggest that regional variation is more related to the supply of orthopedic and neurosurgeons<sup>3</sup> as well as the "surgical signature" phenomenon, which is idiosyncratic to a particular region.<sup>4,5</sup> Furthermore, the rate of lumbar surgery in the United States is double the rate in other developed countries such as Western Europe, Australia, New Zealand, and Canada.<sup>6</sup> It remains unclear if more surgery means better patient outcomes.

It would be hard to dispute the direct relationship between the proliferation of spinal fellowships and hence the number of spine surgeons with the increased number of spinal procedures being performed. Shortly after I be-

From the Department of Orthopedic Surgery, University of California, San Francisco, San Francisco, CA.

The manuscript submitted does not contain information about medical device(s)/drug(s).

No funds were received in support of this work. No benefits in any form have been or will be received from a commercial party related directly or indirectly to the subject of this manuscript.

Address correspondence to David S. Bradford, MD, Department of Orthopedic Surgery, University of California, San Francisco, 500 Parnassus Avenue, MU320-W, San Francisco, CA 94143-0728; E-mail: [bradford@orthosurg.ucsf.edu](mailto:bradford@orthosurg.ucsf.edu)

gan my practice with Dr. Moe in 1970, he initiated what was then the only spine fellowship in North America. Our first fellows were Dr. Ed Dawson and Dr. Claudio Pedras. Now there are over 85 fellowships listed by the North American Spine Society. It is true that many of these fellowships go unfilled, but it is also true that many are training 2 and 3 fellows per year. In Orthopedic Surgery in 2000, the numbers of Part II applicants for the American Board of Orthopedic Surgery certification, which declare to be examined as spine surgeons, were 51; in 2004, it was 84. Neurosurgical data suggest that approximately 130 residents complete their training each year.<sup>7</sup> All are trained in basic spine surgery and many in more complex spinal reconstruction. One might reasonably ask: are we turning out the best possible product and fulfilling a society need? It is difficult to imagine that this "need" is not porous or artificial. If one is "trained to operate and has become specialized only in technical and procedural skills rather than disease management," surgery is what one does. One could not help but agree with the conclusions of Richard Deyo and Alf Nachemson's recent article in the *New England Journal of Medicine* titled "Spinal fusion surgery: the case for restraint."<sup>1</sup> We might suggest that the adage of "don't just do something, stand there" might be a more appropriate recommendation to a surgeon when faced with a patient with low back pain feels that fusion is not necessarily the last resort and is usually the preferred option.

The disturbing lack of evidenced-based approach to managing patients with back pain is not dissimilar to our treatments of other conditions. Historically, surgeons and physicians have often touted procedures that ultimately proved to be disappointing.<sup>8</sup> Internal mammary ligation for angina, gastric freezing for ulcers, radical mastectomy for breast cancer, and the IDET procedure for managing discogenic back pain are just a few. In fact, as Cherkin *et al*<sup>3</sup> noted in a recent article in *Arthritis and Rheumatism*, "Who you see as a patient for low back pain, determines what you get." Neurosurgeons and neurologists are more likely to order imaging studies, physiatrist and neurologist, EMGs, and rheumatologist, lab tests. The cascading effect of new medical technology referred to by Mold and Stein in a classic article in the *New England Journal of Medicine* in 1986<sup>9</sup> may lead to unanticipated adverse effects, *i.e.*, coronary angiography in low-risk patients leading to unnecessary invasive coronary interventions, electronic fetal monitoring leading to unnecessary C-sections, and spinal MRIs/discograms leading to unnecessary spine surgery.<sup>10</sup>

Financial incentives for surgeons, hospitals, and the device industry no doubt also play a role. The impressive growth and profits that have come to the device industry over the last few years have made this sector the darling of Wall Street.<sup>2</sup> In fact, Spine and Cardiac top the device industry in growth prospects and opportunities. Therefore, it is indeed understandable that industry has taken such a leadership and proactive role in introducing new technology, initiating and leading outcome studies, facil-

itating paper publication, including their own statistical analysis, and organizing and financing CME courses.<sup>11</sup> The medical industrial complex or interaction has often proven helpful and symbiotic with the surgical and medical specialties, and has facilitated great advances and breakthroughs for our patients. It has and does, however, come with a price and runs the risk of scientific opacity and loss of objectivity, if not the potential for ethical conflicts.<sup>12</sup> The recent passage of the Sabanes-Oxley Bill will likely bring some clarity and hopefully prove beneficial to all parties in the long run. A better understanding of the optimal way to introduce new technology as outlined by Malchau<sup>13</sup> will prove useful. Finally, we as providers must remember that the health pie is only so large. If technology advances (for example, BMP, disc replacement, *etc.*) are not matched by cost savings from decreases in revision rates, decreased use of iliac bone grafting, shortened length of stay, *etc.*, then reimbursements from the insurance industry and government are likely to be cut further, and I might add, first to the provider, perhaps to the device industry, and I suspect lastly to the hospital. It is expedient that we as a research society become knowledgeable in healthcare cost assessment, and as Bozic *et al*<sup>14</sup> pointed out in a recent article in the *Journal of Bone and Joint Surgery*, policy makers and payers will look to us as clinicians for leadership and scientific evidence to support the efficacy and cost-effectiveness of new technologies. This is an opportunity that we must not pass by.

Finally, the complications from repeat surgery, instrumentation failure, infection, chronic pain, paralysis, and cardiopulmonary failure are well known and in fact may leave the patients far worse off after the revision than before the index procedure. To be fair and put the issues in perspective, the proliferation of spinal surgery and the controversy centering around fusions are those primarily being done for back pain (75%) and less frequently deformity, infection, spondylolisthesis, fractures, *etc.* (25%), yet many of us in AMCs as well as community-based practices doing primary tertiary referral work cannot help but be struck that the rapidity of many of our spinal colleagues to rush to cure degenerative disc disease with the knife before a reasonable trial of conservative treatment, or to rush to do a combined procedure on an adult patient with a 60° curvature, who has only mild back discomfort, and normal pulmonary function, with no evidence of progression and to stress to the patient with almost religious zeal that surgery is required in order to prevent cardiac pulmonary failure and early death! Surely we can do better, and as professionals we must profess to put the patients best interest ahead of our own.

### ■ Issues Facing Academic Spine Surgery

The issues facing academic spine surgery at academic medical centers are likewise substantial if not alarming. One might pick up any recent orthopedic or neurologic surgical journal and read the following "wanted . . . Ac-



ademic Orthopedic/Neurosurgical spine surgeon—at the Assistant or Associate Professor rank, with specialized training in spine surgery. Candidates for the position must be board eligible, or board qualified, committed to teaching and clinical care and basic research. Salary commensurate with academic qualifications.”

This triple threat, of course, no longer exists (if it ever did) and more disturbing I feel is the ever increasing numbers of well-trained, technically competent surgeons, but the steady decline of surgeons capable of doing basic, investigative work or clinically important outcomes studies. The phrase “the Clinical Investigator as an Endangered Species” was first coined by James Wyngaarden in his presidential address to the Association of American Physicians on May 6, 1979.<sup>15</sup> The steady decline of the physician-investigator or physician-scientist has been noted by numerous authors since, and it has occurred across the board of all specialties.<sup>16,17</sup> Our specialty of Orthopedic surgery has become the most endangered of all, I feel, because we were the Cinderella of medicine for so long. For instance, in 1994, we received as a specialty only 0.3% of the total NIH budget, and by the year 2003 we received approximately 0.07% of the total.<sup>18</sup> Although the precise numbers are hard to determine, as Hurwitz and Buchwalter<sup>19</sup> and others have pointed out, the numbers of PIs and Co-PIs that have sustained productive research programs for over a decade has fallen sharply. Looking specifically at NIAMS data for 2004, of the 50 grants awarded for Spine, greater than one half were awarded to nonorthopedic departments. Of the remaining, I could identify only 5 Orthopedic Surgeon PIs, of which only 4 are spine surgeons. Looking at the list of the grants awarded for study of the Intervertebral Disc, of the 55 grants awarded, I could determine for certain that 5 were awarded to Orthopedic Surgeons as the PIs. Certainly, many of the grants have Orthopedic Surgeons as Co-PIs, yet the numbers are disturbing nonetheless. Surely, we continue as a specialty to attract the best and brightest medical students from our leading universities. Many of our applicants hold combined MD/PhD degrees, or have had substantial research experience, and are in the top 5% of their class, yet an investigative career is usually not an election or considered an option for the vast majority. We all recognize the importance of the clinician-scientist that is capable of doing disease and patient-oriented research, at the basic level, translational level, or purely clinical level. As Dr. Doug Jackson has stated, they pose the relevant clinical questions that form the basis of specific investigations leading to clinically relevant, effective, and safe solutions.<sup>20</sup> They are also often our most valued teachers and leaders and raise the bar for our specialty. They are indeed the future. . . . Why are their numbers decreasing?

The reasons are multifactorial. As Drs. Buckwalter and Hurwitz have noted (and I agree), when junior faculty embark on their academic career, they may quickly learn that the path to promotion lies in publications. The

most expeditious and quickest path to build up a CV is by latching onto the clinical work that others in the department have already done, and carry out retrospective reviews and case reports from these data. In terms of significant contribution to orthopedic science, this type of activity often proves trivial, and is not sufficient for producing a well thought out project that leads to grant support. It also muddies up our journals and leads to excessive noise at our scientific meetings.

On the other hand, if the individual has been well trained in science and holds an MD/PhD, they will find by the time that they have finished their residency that they may have difficulty staying current with the field, or their scientific interest may have waned and full-time clinical practice is more appealing.

The other obstacle is certainly financial. With reimbursements ever decreasing, and departmental overhead barely sustainable, the need for new faculty members to “carry their weight” is often explicitly stated. This may prove particularly difficult for those trained as spine surgeons (and I might add sports medicine and hand) as these practices even those in AMC may generate substantial dollars, and portions of their revenue are often necessary to support other members of the department. For the spine surgeon, the differences in earnings between academic and private practice may prove too great an enticement. These facts, coupled with the debt that most residents and fellows members have accumulated by the time they have started practice, may make the most focused clinician-scientist reassess career options.<sup>16,21</sup>

Folkman has noted that two other factors may play a role: “the reverse discrimination syndrome” and the “surgical personality.”<sup>22</sup> The former is applied to clinician-scientist that is looked down on by their colleagues as less than clinically competent because of their scientific bent. I personally have seen this only rarely and feel that perhaps it is more institutionally than departmentally related. At UCSF, I am happy to say that there is more subtle discrimination if one is NOT doing research. The other impediment perhaps could be considered the surgical personality. . . . (I call it also the Surgical Adult Attention Deficit Disorder.) Dr. Folkman has described it as the obsessive need to be “busy” and needed by patients as well as the certainty of conviction or “I may be wrong, but never in doubt.” In the clinical setting, this may be reassuring to a patient with an emergency life-threatening complication (rarely otherwise), but in the laboratory setting, where reflection and cognition as well as healthy skepticism are the pathway and hallmark of success, this self-assuredness may be the pathway to failure.

### ■ Solutions and Opportunities

Solutions to resolve the issues outlined for the specialty of spine as well as its practitioners and researchers are not mutually exclusive. They revolve around changing the culture, enhancing mentorship, fund raising, education (medical students, residents, and fellows), nurturing junior faculty, grantsmanship, and collaborative re-



search and investigative interaction between clinician scientist and clinical practitioners.<sup>16,23-28</sup>

An academic culture that neither encourages nor rewards research, but rather sacrifices it on the altar of departmental efficiency and financial gain is indeed one that sows the seeds of its demise. The leadership and chairs of AMCs must place a top priority on the nurturing and support of the clinician-scientists. Recently, at the AOA/AOS meeting in Boston (June 23, 2004), Regis O'Keefe took a poll of the Orthopedic chairs and program directors, asking: 1) Did they feel that clinical research was important (90% agreed). 2) Did they feel that Basic research in an ortho department was important (only 50% agreed) 3) Did they have a fund raising program (only 66% did). Clearly, if the orthopedic chairs and program directors do not support research, other than retrospective reviews and case reports, we are in dangerous waters. If residents go through 5 years of training, never needing to question a procedure, critically review a scientific paper, or a presentation, and do not develop the analytical skills to do so, or to understand evidence-based approaches to medicine, statistics, and cost/benefit analysis, our specialty and patients will pay the price.

Mentorship is critical to any successful training program. This needs to start at the medical student level. In many programs, we are disadvantaged in that surgical specialties are given little time in the medical school curriculum to teach students. But for those programs that can successfully access their students, show an interest and successfully mentor them, get them involved in a clinical study, or work with members of the department in the research lab for a year or two, they will find that their resident pool will become more exceptional and grow and advantage the department as well as our specialty. For those interested students, the Medical Scientist Training Programs for combined MD/PhD pathway is also an available opportunity.<sup>17,29,30</sup>

To provide financial resources for quality resident education, startup funds for junior faculty, laboratory support not covered by grants, and dollars for educational enrichment for the faculty, an active philanthropic mission is essential. No longer can one rely on clinical income alone to fund these activities. There are just not enough dollars in the system. We are advantaged as orthopedic surgeons and specifically spine surgeons in that our subspecialty does indeed significantly improve the quality of life of our patients, and we are well reimbursed for our efforts. Our patients are in fact often as grateful for the relief of their pain and improvement in function as those patients following a successful arthroplasty. A focused effort at fundraising with a clear definition of goals will enhance our specialty and improve our training programs.

The resident years are an excellent period to direct those interested and qualified graduates toward a clinician-scientist career. A research year or two during their residency may prove very valuable. This may be difficult

for small programs or ones that do not have the necessary resources, but for programs that do, we have found this quite beneficial. When I came to UCSF in 1990, we had no significant research program or facilities. Our immediate goal was to build a nucleus of the best possible basic science group of PhDs that were jointly recruited with the basic science community, then provide them with the best possible resources and start up funds to do the job. Their only charge requirement was a commitment to interact with interested faculty, be willing to teach our residents, and collaborate with those interested and able. After several years of building our facilities, recruitment, nurturing, and supporting them (along with the necessary fund raising), we now have five well-funded research groups at three campuses. A cell/molecular biology and bioengineering group at UCSF campus, a cell/molecular biology group at San Francisco General Hospital, and a cell biology group at the VA are all run by full-time faculty members. The facilities are run by PhDs and clinician-scientist MDs that have grown with the program over the past 10 years. We have also made a commitment by the department to have one interested and committed resident to spend their third year in the research labs. We understand this will not necessarily produce a grant-supported molecular biologist, but rather an inquisitive surgeon, that understands research and can develop collaborations with the full-time scientist placing them in a position to be a Co-PI on NIH grants, and eventually a PI. So far this has exceeded our expectations. Of the 6 that have completed their residency and spent a year in the labs, 5 have gone on to full-time academics, one part-time. Four of the 6 are fellowship-trained spine surgeons, two of these are competitively grant funded. Three of our residents have been MD/PhDs, and all are now actively involved in research and clinical practice. Support for this effort has come from the Orthopedic Research and Education Foundation, our lab PIs mentors, departmental research funds from donations, and faculty commitments from practice funds. This approach will not work for every group, but with a departmental commitment and scientific resources, it is possible.

Fellowships present a challenge to our specialty and to those that run the fellowships. The diversity of experience, the lack of standards or requirements for establishing a fellowship, and the variability of director commitment to education have turned many of these experiences into what I feel might be appropriately referred to as little more than "*in vivo* motor skills" labs. What is more troubling is the continued expansion of fellowships, now numbering in excess of 85. For those programs and surgeons offering fellowships, it is expedient to critically analyze the rationale for their existence. We need to ask ourselves: are our fellows merely being used as cheap sources of labor allowing our program to pay for their efforts by charging assistant fees, and increasing the amount of surgery we can do in the available

time, are our fellowships merely status symbols and marketing tools, or is our fellowship designed to truly educate the fellow in the overall management of spinal disease and disability? Surgical volume should not be a rationale for a fellowship or the rationale for expanding an existing one. In these cases, it may be far better for the fellowship director and the fellows education experience to substitute physician assistants if additional help is needed. Fellowships that involve only a single surgeon mentor, or ones that the fellow must obtain additional "experience" at another institution during the year for completion of the education objectives, would lead one to question its validity. NASS has put together fellowship guidelines that all fellowship directors should periodically review. I would suspect that we as a group are falling far short of these stated goals.

Can we assure that some of the fellowship-trained spine surgeons will become future clinician-scientists, thereby improving science that our specialty sorely needs? Several years ago, we initiated a program while I was at Minnesota that we have continued at UCSF, that is, having two fellowship tracts: one a clinical 1-year program, the other a 2-year program, with the first year in the orthopedic research labs working with basic scientist and one of the spine surgeons. The assigned project would dovetail with work we were doing in the lab so the learning curve would not prove too steep. Of the 7 fellows that have participated in this 2-year program, 4 are in AMC teaching programs and are moving forward in their careers as clinician-scientists. Again, this will not work for all centers, but for those with the facilities and mentorship, we would recommend it. Furthermore, for those interested in an academic career, development of a postgraduate educational program with tutorials on how to plan, conduct, and communicate clinical research would be an important step. Such a program has been done in Brazil and includes: 1) planning a clinical study, 2) the types of clinical trials, and emerging issues regarding how they are used, 3) managing data during a clinical trial, 4) statistical issues in clinical research, 5) ethics of human experimentation, 6) publishing or presenting clinical data, 7) special issues in medical writing, and 8) cost-benefit analysis and what it means, *i.e.*, how new technology should be evaluated.<sup>31</sup> Some institutions already have this as part of their K30 (clinical research curriculum) awards, which are used to develop a mentor-supervised program in clinical research. If a department or fellowship program does not have this opportunity available, each director of the fellowship should consider working with their institutions and initiate an institutionally wide tutorial for all fellows of all specialties and provide time and support for this important experience. By this activity, I am certain that our research as a spine society will prove less trivial, more evidenced based, and more exciting.

The problems that a junior faculty member faces in becoming an established clinician-scientist are well known. As outlined by Goldstein and Brown,<sup>17</sup> three

paths are open to the research-oriented MD who has finished his or her training: pure basic research, disease-oriented research, and patient-oriented research. Pure basic research is a possible but an unlikely path for the trained orthopedic spine surgeon.

Disease-oriented research is enticing and possible for the trained spine surgeon. The clinician scientists may start on this track with a patient problem and then gravitate to pure basic science (it is of interest that the last three Nobel prize winners at UCSF tracked this path) or continue to split their activities between clinical care and research. For instance, 3 days in clinical care and 2 days in the lab would seem reasonable. A basic amount of knowledge such as some previous work in a research lab is essential. To be competitive for NIH support immediately at the completion of training requires at least a 2- to 3-year commitment in basic science under an established mentor or significant experience in a research lab in advance. Those that wish to take this route will find that there are funds available and, in the proper environment with supportive mentorship, they will be successful. For basic science, K08 awards requiring 75% of a 40-hour workweek are available for 3 to 5 years. Mentor involvement and collaboration are indeed helpful, and mentors can be external to the university. This external mentorship also applies to RO1 and R03 and well as program project grants.

Finally, those engaged in patient-oriented research may develop a very promising and successful career by collaboration with basic scientists likewise interested in translational research. Both bring something to the table; and if the surgeon becomes actively involved and provides intellectual capital to the problem, assist with grant writing, becoming a Co-PI, a very successful and mutually beneficial outcome is likely. Furthermore, other options in this pathway include clinical research awards, such as the K23 and K24 (mid-career) award for those programs that have a clinical research centers. Additional opportunities for external funding include the OREF (which last year awarded \$3.4 million for research and raised an additional \$1.9 million for use by specialty societies. NASS and the SRS are also opportunities. The SRS has now raised in excess of \$2.6 million. I would be less than candid if I did not note that of the 400 plus members of the SRS only 35 are Shands level donors, a donation of only \$20,000 (that can be done over several years). This amount is less than the reimbursement to the surgeon for one combined deformity procedure! I was also surprised to note that only \$40,000 has been pledged by the SRS Board to research this past year and next year. I would ask the Board of Directors, in light of the total dollars in the endowment, to review their granting policy and hence the amount of funds being awarded yearly toward grants. If projects are worthy, the endowment investment gain would find best use in supporting research rather than further "growing" the endowment. I would submit that the rainy day is now and not several years down the road! I would also recommend that the

Board of Directors consider using these funds for research and educational activities of those spine surgeons beginning and committed to a physician-scientist career. At the same time, I would strongly encourage our membership to step up to the plate and join the Shands circle. You are the leaders of the spine field! Your actions will speak louder than any paper you might present. Nothing could be more important to our membership and the future of our profession.

Certainly none of these paths will be appropriate to everyone throughout their surgical career. Attrition will occur, and fine academicians will depart to other institutions or to clinical practice without "academic" involvement. But this eventuality is not the end of the world. Faculty turnover can be an opportunity. The tree can be pruned and new, more vibrant growth possible. But for those that stay the course and pursue a scientific approach to clinical practice, either at an AMC or a pure clinical setting, they will find a career and life that is filled with excitement, challenges, satisfaction, and never ending opportunities.

Initiating these efforts cannot only strengthen academic spine surgery but also those practices removed from academic medical centers. Not only will the applicant pool of better-educated residents be entering practice and fellowships, but improved rationale and outcomes for what we do will be the likely result. We will have the knowledge base to more carefully evaluate emerging technology and be more critical of newness for newness sake. We, as a community of spine surgeons, will find that multicentered trials where community and academic surgeons pool their resources will be more successful and more fundable as we all have a better understanding of how to do outcomes research—the SPORT trial that Dr. Weinstein has initiated is a stellar example of what can be accomplished—and more fully understand healthcare technology assessment and cost-benefit analysis. For instance, our efforts as a group might better be directed on who gets a fusion, rather than how to do it. This concept could also apply to innovative displacement technologies, such as disc and nucleus replacement, dynamic stabilization, or interspinous distraction procedures (the X-stop/Wallis implants).

In short, we as outcomes specialists, evidenced-based practitioners, and clinician-scientists will be the determiners of what is appropriate for our patients rather than the third party carriers or device companies whose mission may not necessarily be in parallel with ours. By these activities and our commitment to evidenced-based outcomes and sound research, we can then direct our specialty to what we inherently know is best and correct for our patients without being swayed by the seductiveness of instant gain and unproven technology. This is the definition and mark of a true professional and what we as a society of academicians must continue to profess.

Thank you for the privilege of being your 2004 Harrington Guest Lecturer.

## References

1. Deyo RA, Nachemson A, Mirza SK. Spinal-fusion surgery: the case for restraint. *N Engl J Med* 2004;350:722–726.
2. Mendenhall S. Spinal surgery update. *Orthop Network News* 2003;14:1–3.
3. Cherklin DC, Deyo RA, Wheeler K, et al. Physician variation in diagnostic testing for low back pain. *Arthritis Rheum* 1994;37:15–22.
4. Weinstein J, Bronner K, Morgan T, et al. Trends and geographic variations in major surgery for degenerative diseases of the hip, knee and spine. *Health Aff* 2004 (Web exclusive, <http://content.healthaffairs.org/cgi/content/abstract/hlthaff.var.81v1?ck=nck>).
5. Weinstein JN. In: Birkmeyer J, Wennberg JE, Cooper M, eds. *The Dartmouth Atlas of Musculoskeletal Health Care*. Hanover, NH: NHL Center for the Evaluative Clinical Sciences, Dartmouth Medical School, 2000.
6. Hanley EN, David SN. Lumbar arthrodesis for the treatment of back pain. *J Bone Joint Surg Am* 1999;81:716–730.
7. Gottfried OM, Rovit RL, Couldwell WT. Trends in the neurosurgical workforce in the United States (unpublished).
8. Groopman J. A knife in the back. *The New Yorker*; April 2002;66–73.
9. Mold JW, Stein HF. The cascade effect in the clinical case of patients. *N Engl J Med* 1986;314:512–514.
10. Deyo R. Cascade effects of medical technology. *Annu Rev Public Health* 2002;23:23–44.
11. Burkus JK, Heim SE, Gornet MF, et al. Is INFUSE bone graft superior to autograft bone? An integrated analysis of clinical trials using the LT-CAGE lumbar tapered fusion device. *J Spinal Disord Tech* 2003;16:113–122.
12. Mirza SK. Accountability of the accused: facing public perceptions about financial conflicts of interests in spine surgery. *Spine J* 2004;4:491–494.
13. Malchau H. Introducing new technology: a stepwise algorithm. *Spine* 2000;25:285.
14. Bozic KJ, Pierce RG, Herndon JH. Health care technology assessment: basic principles and clinical applications. *J Bone Joint Surg Am* 2004;86:1305–1313.
15. Wyngaarden JB. The clinical investigator as an endangered species. *N Engl J Med* 1979;301:1254–1259.
16. Byrne E. The physician scientist: an endangered breed? [Editorial] *Intern Med J* 2004;34:75.
17. Goldstein JL, Brown MS. The clinical investigator: bewitched, bothered, and bewildered—but still beloved. *J Clin Invest* 1997;99:2903–2912.
18. Thompson RC. The future of academic medicine or is academic medicine to be determined by technology? *Iowa Orthop J* 2002;22:116–119.
19. Hurwitz SR, Buckwalter JA. The orthopaedic surgeon scientist: an endangered species? *J Orthop Res* 1999;17:155–156.
20. Jackson DW. The orthopaedic forum: the orthopaedic clinical-scientist. *J Bone Joint Surg Am* 2001;83:131–134.
21. Ahn J, Jackson DW. Encouraging the development of the orthopaedic clinician-scientist. *J Bone Joint Surg Am* 2002;84:878–879.
22. Folkman J. Surgical research: a contradiction in terms. *J Surg Res* 1984;36:294–299.
23. Henke PK. Collaboration based on the clinical model can help grow the physician-scientist. *Am J Surg* 2003;185:42–44.
24. Zemlo TR, Garrison HH, Partridge NC, et al. The physician-scientist: career issues and challenges at the year 2000. *FASEB J* 2000;14:221–230.
25. Rosenberg LE. The physician-scientist: an essential—and fragile—link in the medical research chain. *J Clin Invest* 1999;103:1621–1626.
26. Rosenberg LE. Physician-scientists: endangered and essential. *Science* 1999;283:331–332.
27. Faxon DP. The chain of scientific discovery: the critical role of the physician-scientist. *Circulation* 2002;105:1857–1860.
28. Sargent MC, Sotile W, Sotile MO, et al. Topics in training: stress and coping among orthopaedic surgery residents and faculty. *J Bone Joint Surg Am* 2004;86:1579–1586.
29. Clark JM, Hanel DP. The contribution of MD-PhD training to academic orthopaedic facilities. *J Orthop Res* 2001;19:505–510.
30. Varki A, Rosenberg LE. Emerging opportunities and career paths for the young physician-scientist. *Nat Med* 2002;8:437–438.
31. Badaro R, Jones TC. Preventive healthcare for an endangered species—the physician scientist. *Braz J Infect Dis* 2001;5:101–3.



# A Multicenter, Prospective, Randomized Trial Evaluating the X STOP Interspinous Process Decompression System for the Treatment of Neurogenic Intermittent Claudication

## Two-Year Follow-Up Results

James F. Zucherman, MD,\* Ken Y. Hsu, MD,\* Charles A. Hartjen, MD,†  
Thomas F. Mehalic, MD,‡ Dante A. Implicito, MD,§§ Michael J. Martin, MD,¶  
Donald R. Johnson II, MD,|| Grant A. Skidmore, MD,\*\* Paul P. Vessa, MD,††  
James W. Dwyer, MD,†† Stephen T. Puccio, MD,§§§ Joseph C. Cauthen MD,¶¶ and  
Richard M. Ozuna, MD||||

**Study Design.** A randomized, controlled, prospective multicenter trial comparing the outcomes of neurogenic intermittent claudication (NIC) patients treated with the interspinous process decompression system (X STOP) with patients treated nonoperatively.

**Objective.** To determine the safety and efficacy of the X STOP interspinous implant.

**Summary of Background Data.** Patients suffering from NIC secondary to lumbar spinal stenosis have been limited to a choice between nonoperative therapies and decompressive surgical procedures, with or without fusion. The X STOP was developed to provide an alternative therapeutic treatment.

**Methods.** 191 patients were treated, 100 in the X STOP group and 91 in the control group. The primary outcomes measure was the Zurich Claudication Questionnaire, a patient-completed, validated instrument for NIC.

**Results.** At every follow-up visit, X STOP patients had significantly better outcomes in each domain of the Zurich Claudication Questionnaire. At 2 years, the X STOP patients improved by 45.4% over the mean baseline Symptom Severity score compared with 7.4% in the control group; the mean improvement in the Physical Function domain was 44.3% in the X STOP group and –0.4% in

the control group. In the X STOP group, 73.1% patients were satisfied with their treatment compared with 35.9% of control patients.

**Conclusions.** The X STOP provides a conservative yet effective treatment for patients suffering from lumbar spinal stenosis. In the continuum of treatment options, the X STOP offers an attractive alternative to both conservative care and decompressive surgery.

**Key words:** prospective randomized study design, lumbar spinal stenosis, neurogenic intermittent claudication, epidural injection, laminectomy, interspinous process decompression. **Spine 2005;30:1351–1358**

Studies evaluating neurogenic intermittent claudication (NIC) secondary to lumbar spinal stenosis (LSS) indicate that 3 to 4% of patients with low back pain who see a general physician have LSS, and 13 to 14% of patients with low back pain who see a specialist have LSS.<sup>1–4</sup> The cost to society of NIC resulting from medical care, loss of productive work hours, legal costs, and compensation costs is in the tens of billions of dollars in the United States annually.<sup>5,6</sup> The definition, etiology, clinical symptoms, incidence, and treatment of NIC have been well documented and are generally attributed to neural compression at one or more lumbar motion segments.<sup>7–22</sup>

The characteristic symptoms of NIC such as back and leg pain, tingling, numbness, and weakness are generally present depending on the patient's posture, with symptoms exacerbated in positions of lumbar extension such as standing and walking, and relieved in positions of flexion such as sitting or bending forward.<sup>8,12,17,19–21,23–25</sup> The primary level affected is L4–L5, followed by L3–L4, L5–S1, L2–L3, and L1–L2.<sup>14,15,17,22,26</sup> Patients with stable symptoms are treated with a regimen of nonoperative therapy that may include epidural steroid injections, oral steroids, nonsteroidal anti-inflammatory medication, analgesics, physical therapy, and spinal manipulation. The only treatment option available to patients who fail to respond to these therapies is decompressive surgery, such as a laminectomy, which may be accompanied by a fusion. The success rate of decompressive surgery as re-

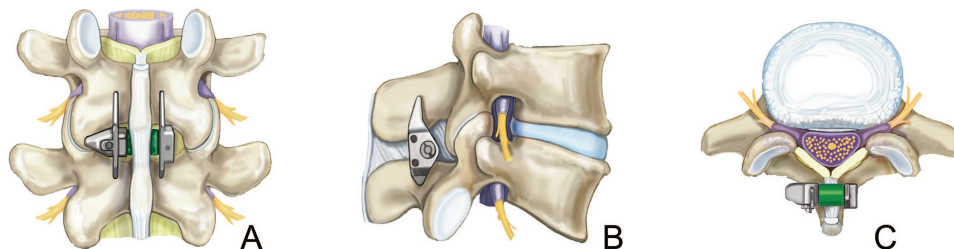
From \*St Mary's Hospital, San Francisco, California; the †Greater Baltimore Medical Center, Baltimore, Maryland; the ‡Maine Medical Center, Portland, Maine; the §Hackensack University Medical Center, Paramus, New Jersey; the ¶Tacoma General Hospital, Tacoma, Washington; the || East Cooper Medical Center, Mount Pleasant, South Carolina; the \*\*East Virginia Medical School, Norfolk, Virginia; the ††Somerset Medical Center, Somerset, New Jersey; the §§Orthopaedic Specialists of Bethlehem, Bethlehem, Pennsylvania; the ¶¶North Florida Regional Medical Center Gainesville, Florida; and the ||||Brigham and Women's Hospital, Boston, Massachusetts.

Acknowledgment date: February 24, 2004. First revision date: June 9, 2004. Second revision date: July 20, 2004. Acceptance date: August 3, 2004.

The device(s)/drug(s) that is/are the subject of this manuscript is/are being evaluated as part of an ongoing FDA-approved investigational protocol (IDE) or corresponding national protocol for the X STOP. Corporate/Industry funds were received in support of this work. One or more of the author(s) has/have received or will receive benefits for personal or professional use from a commercial party related directly or indirectly to the subject of this manuscript: e.g., royalties, stocks, stock options, decision making position.

Address correspondence and reprint requests to James F. Zucherman, MD, St. Mary's Spine Center, One Shrader Street, Suite 450, San Francisco, CA 94117; E-mail: yaquai@aol.com

Figure 1. A schematic of the X STOP *in situ*. The posterior (A), lateral (B), and axial (C) views show that the implant is placed between the spinous process. The lateral wings prevent anterior and lateral migration, and the supraspinous ligament prevents posterior migration.



ported in the clinical literature is quite variable, and the procedure is associated with a relatively high complication and reoperation rate.<sup>27–32</sup> Turner's meta-analysis of 74 published studies of surgery for lumbar spinal stenosis found good to excellent results ranging from 26 to 100%.<sup>33</sup>

The interspinous process decompression system (X STOP) provides an alternative therapy to conservative treatment and decompressive surgery for patients suffering from NIC.<sup>34</sup> The X STOP is implanted between the spinous processes and reduces pathologic extension at the symptomatic level(s), while allowing flexion and unrestricted axial rotation and lateral bending.<sup>35</sup> Biomechanical studies have shown that the implant significantly reduces intradiscal pressure and facet load and prevents narrowing of the spinal canal and neural foramina.<sup>36–38</sup> The current study reports the 2-year outcomes from a prospective, randomized, multicenter study of NIC patients. The specific aims of the study were to measure the percentage of improvement of patients treated nonoperatively and with the X STOP. The authors hypothesized that the X STOP would be significantly more effective than conservative care at all follow-up visits.

## ■ Patients and Methods

**Patient Selection.** One hundred ninety-one patients were enrolled in a prospective, randomized, controlled trial at 9 US centers over a 15-month period from May 2000 to July 2001. The study was conducted under a Food and Drug Administration-approved Investigational Device Exemption and was approved by the Institutional Review Board at each participating institution before initiation. All patients signed an Institutional Review Board-approved informed consent form before participation in the study. Patient eligibility to participate in the study was based on the following key inclusion and exclusion criteria.

**Key Inclusion Criteria.** Patients had to be at least 50 years old and have leg, buttock, or groin pain with or without back pain that was relieved during flexion. To identify a study population of patients with more moderate symptoms of NIC, patients had to be able to walk at least 50 feet.

**Key Exclusion Criteria.** Patients could not have a fixed motor deficit, cauda-equina syndrome, previous lumbar surgery of the stenotic level, or spondylolisthesis greater than grade I on a scale of I to IV at the affected level(s).

**Randomization.** Block randomization by site was used to ensure a balanced proportion of X STOP and control subjects

in each clinical site and for the entire study. The date of surgery was considered as the treatment date for X STOP patients, and the date of the initial epidural injection was considered as the treatment date for control patients.

**Control Group.** Nonoperative therapy was selected as a control in the current study, both because it is the standard of care in the treatment for patients with mild to moderate NIC and because implantation of the X STOP, like nonoperative care, does not require the patient to undergo a highly invasive procedure. Patients randomized to the control group received at least one epidural steroid injection following enrollment and were prescribed additional epidural steroid injections, nonsteroidal anti-inflammatory medications, analgesics, and physical therapy as necessary. Physical therapy consisted of back school and methods such as ice packs, heat packs, massage, stabilization exercises, and pool therapy.

**X STOP Group.** Patients randomized to the X STOP group underwent surgery for implantation of the device, which consists of two components: a spacer assembly and a wing assembly. The X STOP is placed between the spinous processes from a lateral direction without resecting the supraspinous ligament or the removing of any tissue (Figure 1). The surgical technique is described in more detail by Zucherman *et al*.<sup>34</sup>

**Outcomes Assessment.** Assessments were made before treatment (baseline) and at 6 weeks, 6 months, 1 year, and 2 years following treatment. Assessment of the primary outcome was based on data collected using the patient-completed Zurich Claudication Questionnaire (ZCQ), which consists of Symptom Severity and Physical Function domains that are completed before and after surgery and the Patient Satisfaction domain that is completed after surgery.<sup>39,40</sup> The mean percent improvement from baseline in the Symptom Severity and Physical Function domains was calculated for each patient at each time point. Also, the proportion of patients in both groups who were clinically significantly improved and who were satisfied with their treatment was compared at each follow-up time point. The mean percent improvement from baseline in the Symptom Severity and Physical Function domains was compared between the X STOP and control groups using an ANOVA with a level of significance of 0.05. The percentage of patients who had significant clinical improvement in each domain was compared between the X STOP and control groups using the Fisher exact test with a level of significance of 0.05. Pretreatment variables including baseline scores, patient demographics such as age or gender, the presence of comorbid conditions, and operative variables for X STOP patients were correlated with treatment success using univariate and multivariate regression analyses to determine predictors of outcomes. All independent variables associated with levels of sig-

**Table 1. Patient Demographics and Baseline Variables**

Variable	X STOP	Control	P *
Age (years)	70.0 (9.8)	69.1 (9.9)	0.513
Height (cm)	170.9 (9.7)	168.4 (11.2)	0.117
Weight (kg)	80.4 (15.8)	81.8 (18.9)	0.569
Baseline SS	3.14 (0.56)	3.10 (0.51)	0.582
Baseline PF	2.48 (0.48)	2.48 (0.51)	0.938
Spondylolisthesis present	35/100	24/90	0.272

Note. mean (SD).  
\* Student's *t* test.

nificance  $<0.1$  in the univariate analyses were included in multiple logistic regression models with stepwise selection of variables.

**Radiographic Analysis.** All patients underwent a radiographic examination at each follow-up visit. The examination included anteroposterior and sagittal plain radiographs of the lumbar spine in the neutral or standing position. The distance between the spinous processes of the implanted levels of X STOP was compared between the 6-week and 2-year radiographs using the method of Neumann *et al.*<sup>41</sup> Additional measurements were made to determine if the X STOP resulted in any radiographic changes to the lumbar spine that could be of potential clinical significance, such as whether there was an increase or decrease in the angulation or curvature of the spine or whether there was an increase or decrease in the percentage of spondylolisthesis. Measurements in the X STOP patients were compared with measurements made in control patients at 1- and 2-year follow-up. All measurements were made by an independent radiologist and comparisons were performed using Student's *t* test with a level of significance of 0.05.

**Safety.** Complications were assessed intraoperatively and after surgery until patients completed the study.

## ■ Results

### **Demographics and Baseline Variables**

There were no significant differences in age, height, or weight between the two groups (Table 1). The mean age was 70 years in the X STOP group and 69.1 in the control group. Also, there were no significant differences in baseline Symptom Severity or Physical Function domain scores between the two groups (Table 1). Spondylolisthesis of Grade I or less was present in 35% of the X STOP patients and 27% of the control patients; the remaining patients had no spondylolisthesis present.

### **Operative Details**

A total of 136 levels were implanted in 100 patients; 64 single levels and 36 double levels. The procedure took an average of  $54 \pm 18$  minutes (mean  $\pm$  SD), and the average blood loss was 46 mL. The most common level implanted was L4–L5 (89/136), and the second most common level was L3–L4 (43/136). The procedure was performed under local anesthesia in 97 patients and under general anesthesia in 3 patients. Ninety-six patients were in the hospital less than 24 hours and four stayed greater than 24 hours.

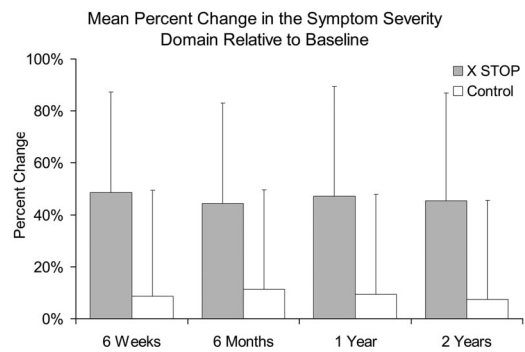


Figure 2. The mean percent change of the Symptom Severity scores relative to each patient's baseline score. At each time point, the mean percent change of the X STOP patient's score was significantly greater than that of the control patient's score ( $P < 0.001$ ). There were no significant differences between time points for either the X STOP ( $P > 0.590$ ) or control groups ( $P > 0.900$ ). The percent change for each patient at each time point was calculated as the change from baseline relative to the baseline score, *i.e.*,  $(\text{Baseline score} - \text{score}_t) / \text{Baseline score}$ .

### **Epidural Injections**

All 91 patients in the control group received an epidural steroid injection following enrollment. An additional 125 injections were administered to control group patients over the course of the study, for a total of 216 injections. Fifty-nine control group patients received at least one additional injection after the initial injection at baseline: 22 patients received 1 additional injection, 21 patients received 2 additional injections, 8 patients received 3 additional injections, and 8 patients received 4 or more injections.

### **Patient Follow-up**

At 2-years follow-up, data from 93 of the 100 X STOP patients and 81 of the 91 control patients were available for analysis. In the X STOP group, seven patients were lost to follow-up; four patients died, two patients failed to complete the ZCQ, and one patient withdrew. In the control group, ten patients were lost to follow-up; three patients died, one patient could not tolerate the initial epidural steroid injection which was aborted, and six patients withdrew. Outcomes from these patients are not included in the results. None of the deaths in the study were attributable to treatment in either group.

### **Primary Outcomes**

The mean percent improvement of the Symptom Severity and Physical Function domain scores in the X STOP group were significantly greater than those of the control group at each time point (Figures 2 and 3). At 2 years, the mean Symptom Severity scores improved by 45.4% from the baseline scores in the X STOP group and by 7.4% in the control group ( $P < 0.001$ ). At the same time point, the mean Physical Function scores improved by 44.3% in the X STOP group and by  $-0.4\%$  in the control group ( $P < 0.001$ ). These findings were consistent for the two domains at all time points (Figures 2 and 3). In the X STOP group, there were no significant differences between the mean percent improvement at any two fol-



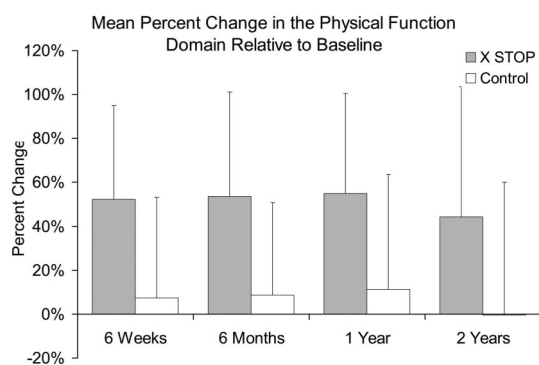


Figure 3. The mean percentage change of the Physical Function scores relative to each patient's baseline score. At each time point, the mean percent change of the X STOP patient's score was significantly greater than that of the control patient's score ( $P < 0.001$ ). There were no significant differences between time points for either the X STOP ( $P > 0.087$ ) or control groups ( $P > 0.270$ ). The percent change for each patient at each time point was calculated as the change from baseline relative to the baseline score, i.e.,  $(\text{Baseline score} - \text{score}_t) / \text{Baseline score}$ .

low-up time points in the Symptom Severity or Physical Function domains ( $P > 0.59$  and  $P > 0.087$ , respectively). In the control group, there were no significant differences between the mean percent improvement at any two follow-up time points in the Symptom Severity or Physical Function domains ( $P > 0.9$  and  $P > 0.27$ , respectively).

At the 2-year evaluation, 56 of 93 patients (60.2%) reported a clinically significant improvement in the Symptom Severity domain compared with 15 of 81 patients (18.5%) in the control group ( $P < 0.001$ ), 53 of 93 patients (57.0%) reported clinically significant improvement in the Physical Function compared with 12 of 81 patients (14.8%) in the control group ( $P < 0.001$ ), and 68 of 93 patients (73.1%) were at least somewhat satisfied compared with 28 of 78 patients (35.9%) in the control group ( $P < 0.001$ ). The proportion of patients who satisfied all three ZCQ criteria was 48.4% in the X STOP group compared with 4.9% in the control group. The percentage of patients with significant clinical improvement at 6 weeks, 6 months, and 1 year has been reported previously.<sup>34</sup>

### Predictors of Outcomes

In the univariate analysis, 13 variables were significantly correlated to patient success in the X STOP group, and three of these variables remained significant in the multivariate model (Table 2). A positive femoral stretch test, the absence of comorbid conditions, and lower surgical blood loss were the most significant predictors of patient success in the univariate analysis and the only significant predictors in the multivariate analysis. No variables associated with the control group were significant in the univariate analysis. The presence of spondylolisthesis was not predictive of outcomes although 55.9% (19 of 34) of the X STOP patients with spondylolisthesis were clinically successful compared with 44.1% (26 of 59) of patients without spondylolisthesis.

Table 2. Predictors of Outcomes

X STOP Group Variable	Univariate		Multi- Factor	
	Estimate	P	Estimate	P
Femoral stretch test	-1.70	0.001 *	-1.50	0.010 †
Comorbid conditions	1.39	0.003 *	1.33	0.013 †
Blood loss	0.02	0.004 *	0.02	0.007 †
ZCQ physical function	-1.40	0.005 *		NS
SF-36 social functioning	0.02	0.010 *		NS
Range of motion-extension	-0.06	0.012 *		NS
SF-36 bodily pain	0.04	0.013 *		NS
Range of motion-rotation	-0.04	0.021 *		NS
Employed	-0.98	0.034 *		NS
Age	0.04	0.048 *		NS
L4-L5 involvement	-2.06	0.058 *		NS
Back pain present	1.14	0.075 *		NS
Use of narcotics	-0.75	0.081 *		NS

NS = not significant.

\* Indicates a level of significance  $< 0.1$ .

† Indicates a level of significance  $< 0.05$ .

### Additional Surgery

Six patients in the X STOP group and 24 patients in the control group underwent decompressive surgery (laminectomy) for unresolved stenosis symptoms during the 2-year follow-up period. Postlaminectomy outcomes are available for 28 patients (6 X STOP and 22 control patients). The mean follow-up time for this group was 12.8 months (range 2.5–26.9 months). The patients undergoing a laminectomy improved by 33.2% in the Symptom Severity domain and by 37.9% in the Physical Function domain. Sixteen of 28 (57.1%) patients had significant clinical improvement in Symptom Severity, 18 of 28 (64.3%) had significant clinical improvement in Physical Function, and 15 of 28 (53.6%) were satisfied with the outcome of their treatment (Table 3). Forty-three percent (12/28) of laminectomy patients met all three of the ZCQ criteria.

### Safety/Complications

No device-related intraoperative complications occurred, and investigators were able to complete implantation of the X STOP in all patients. No procedures were converted to a laminectomy at the time of X STOP surgery.

Three complications occurred intraoperatively or within 72 hours following surgery in the X STOP group (Table 4). There was one episode of respiratory distress and one ischemic coronary episode that resolved without clinical sequelae. One X STOP patient with a history of

Table 3. Comparison of X STOP and Laminectomy ZCQ Outcomes

	X STOP	Laminectomy	P*
Symptom severity	56/93 (60.2%)	16/28 (57.1%)	0.827
Physical function	53/93 (57.0%)	18/28 (64.3%)	0.520
Patient satisfaction	68/93 (73.1%)	15/28 (53.6%)	0.064
Overall success	45/93 (48.4%)	12/28 (42.9%)	0.669

\* Fisher's exact test.

**Table 4. Complications of X STOP and Control Patients**

Complication	X STOP (N = 100)		Control (N = 91)	
Intraoperative or procedure related				
Respiratory distress	1	1.0%	0	0.0%
Coronary episode, ischemic	1	1.0%	0	0.0%
Pulmonary edema	1	1.0%	0	0.0%
Wound dehiscence	1	1.0%	NA	NA
Wound swelling	1	1.0%	NA	NA
Hematoma	1	1.0%	NA	NA
Incisional pain	1	1.0%	NA	NA
Injection intolerance	NA	NA	1	1.1%
Symptom flare	NA	NA	1	1.1%
Leg paresthesia	NA	NA	2	2.2%
Increased back pain	NA	NA	1	1.1%
Heart attack	NA	NA	1	1.1%
Device related				
Malpositioned implant	1	1.0%	NA	NA
Implant dislodgement/migration	1	1.0%	NA	NA
Spinous process fracture	1	1.0%	NA	NA
Increased pain at implant level	1	1.0%	NA	NA

NA = not applicable.

cardiovascular disease developed pulmonary edema 2 days following device implantation. This patient subsequently died. There were four minor operative site-related complications in the immediate postoperative period: one wound dehiscence, one swollen wound that was aspirated, one hematoma, and one report of incisional pain (Table 4). There were three device-related complications in the X STOP group (Table 4). One X STOP patient suffered a fall that caused the implant to dislodge. The dislodged implant was removed without sequelae. An asymptomatic spinous process fracture was diagnosed in another patient on routine 6-month follow-up radiographs, which required no further medical treatment or surgical intervention. One patient reported worsening pain 382 days following treatment, which was determined to be possibly related to the implant. Finally, one implant was placed posterior enough to be considered malpositioned.

Five complications were associated with the epidural injection (Table 4). One patient was unable to tolerate the injection, and the investigator abandoned the procedure; one patient had a severe flare in symptoms and was admitted overnight; two patients had leg paresthesias and were discharged following observation; and one patient sought treatment at an emergency room for back pain 6 hours following the injection. Another patient suffered a heart attack 3 days following treatment; it is unknown whether the heart attack was related to the injection procedure.

Distraction was maintained in 96% of the levels implanted with the X STOP, defined as no measurable change in the distance between the spinous processes when radiographs taken at the 6 week follow-up were compared with radiographs taken at the 2-year follow-up. There were no significant differences between the X STOP and control groups in the mean values of any other radiographic measurements made at either the 1-year or 2-year follow-up visits (Table 5).

## Discussion

Currently available options to treat NIC are limited to either nonoperative therapy or decompressive lumbar laminectomy, with or without a fusion. With the exception of the 1-year report by Zucherman *et al*,<sup>34</sup> no randomized, prospective, multicenter study has been performed on NIC patients to determine the efficacy of either nonoperative therapy or surgical decompression.<sup>33,42</sup> Few studies are prospective, the follow-up data collection methods are unclear, rarely is the data analyzed by someone other than the treating physician, and the outcomes are not assessed at consistent time intervals.<sup>33,42</sup>

The current study reports the 2-year outcomes of NIC patients in a randomized, prospective, multicenter study using a validated, patient-completed instrument to quantify a change in the symptoms, physical function, and

**Table 5. Mean Radiographic Measurements, 12- and 24-Month Follow-up Visits**

Measurement	Follow-up (months)	X STOP	Control	P*
Spinous process Distance (mm)	12	52.1 (7.1)	51.0 (7.0)	0.336
	24	51.8 (7.4)	51.2 (7.1)	0.592
Anterior disc height (mm)	12	9.9 (4.2)	9.7 (3.8)	0.776
	24	9.0 (4.1)	8.9 (4.3)	0.839
Posterior disc height (mm)	12	5.3 (2.5)	5.1 (2.3)	0.626
	24	4.6 (2.3)	4.6 (2.2)	0.935
Treated level angulation (deg)	12	14.6 (7.4)	16.5 (6.7)	0.099
	24	15.1 (7.1)	15.5 (7.6)	0.707
L1–L5 angulation (deg)	12	34.4 (11.9)	33.5 (14.1)	0.701
	24	35.6 (11.5)	32.8 (13.1)	0.198
Foraminal height (mm)	12	23.2 (2.5)	22.5 (2.5)	0.088
	24	21.2 (2.8)	21.5 (2.7)	0.412
Spondylolisthesis (%)	12	4.1 (8.7)	5.9 (9.0)	0.201
	24	4.7 (9.2)	7.0 (10.4)	0.154
L1–L5 coronal curve (deg)	12	4.9 (4.3)	5.8 (5.5)	0.267
	24	6.1 (5.5)	4.9 (4.1)	0.193

Note. Mean (SD).

\* Student's *t* test.

patient satisfaction following treatment for NIC. The results of this study and the previous report by Zucherman *et al*<sup>34</sup> demonstrate that the X STOP significantly improves symptoms and function compared with epidural steroid injections and conservative therapy at 6-week, 6-month, 1-year, and 2-year follow-up.

The presence of comorbid conditions was a negative predictor of outcomes in this trial, which has been noted in outcomes of decompressive surgery for LSS. Katz *et al*<sup>31</sup> reported that patients with greater co-morbidities and worse self-rated health, physical function, symptom severity, and depression were associated with worse outcomes. Jonsson *et al*<sup>30</sup> reported that 23 of the 50 patients had concomitant diseases that affected walking ability and likely the outcomes.

To place the outcomes of the X STOP in the spectrum of current treatment alternatives for NIC, we have compared the outcomes of the relevant literature regarding the safety and efficacy of decompressive laminectomy with the X STOP outcomes.<sup>26,31,32,43–45</sup> In a study by Johnsson *et al*,<sup>45</sup> approximately 60% of the LSS patients treated surgically graded their condition as improved, whereas approximately 40% were either unchanged or worse. In a study by Amundsen *et al*,<sup>26</sup> 15 of 48 (31%) of patients treated surgically assessed their pain as none to light and could be considered clinically successful. In two successive reports by Atlas *et al*,<sup>32,46</sup> between 60 and 70% of the surgical patients were satisfied following surgery, and their predominant symptom was “better” in approximately 55 to 70% of the patients. Using a more rigorous definition of clinical success, Gunzburg *et al*<sup>44</sup> reported that 21 of 36 (58%) reported improvement in three of the four outcomes measures (visual analog pain intensity scale, Oswestry Low Back Pain Disability Questionnaire, Waddell Disability Index, and Low Back Outcome Score), used in their study, and 14 of 36 (39%) reported improvement in all four outcomes measures. Katz *et al* reported outcomes in 197 patients with 2 year follow-up, using the ZCQ and the same success criteria in a patient population similar to those enrolled in this study. Katz *et al*<sup>31,43</sup> reported that 63% of the patients were significantly improved in Symptom Severity, 59% were improved in Physical Function, and 72% were satisfied (Figure 4). Forty-seven percent of patients met all three criteria. These results confirm that outcomes of X STOP patients are comparable with the results of patients undergoing decompressive laminectomy.<sup>26,31,32,43–45</sup> There were also no significant differences in the outcomes of the 6 X STOP and 22 control patients who underwent a decompressive laminectomy compared with the X STOP using the same outcomes measure (Figure 4).

Although the outcomes of the X STOP and surgical decompression procedures are comparable, there are significant differences in the risks associated with the two surgical procedures. Mean operative time for the X STOP procedure was 54 minutes, which is considerably less than the range of 72 to 278 minutes reported for

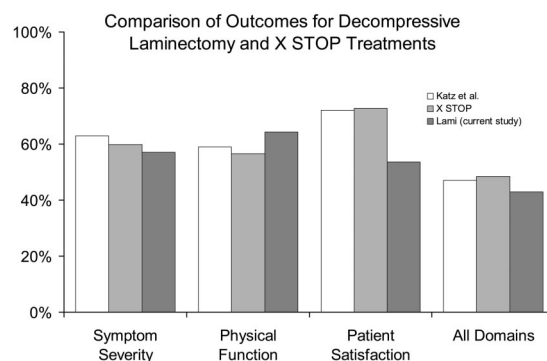


Figure 4. A comparison of outcomes for NIC patients treated with decompressive laminectomy and the X STOP.

laminectomy procedures.<sup>11,30,47–50</sup> Also, the mean blood loss of 46 mL during the X STOP procedure is considerably less than the range of 115 to 1040 mL reported for decompressive surgery.<sup>11,30,47,48,50</sup> Complications reported for laminectomy include paralysis, myocardial infarction, pulmonary embolism, pneumonia, hematoma, deep vein thrombosis, neurologic deficit, deep infection, superficial infection, dural tears, implant failure (when accompanied by a fusion), and pseudarthrosis.<sup>7,32,33,47,51–53</sup> Turner's meta-analysis reported the following complication rates for NIC surgery: perioperative mortality (0.32%), dural tears (5.91%), deep infection (1.08%), superficial infection (2.3%), deep vein thrombosis (2.78%), any complication (12.64%).<sup>33</sup> None of these major complications was reported as a result of the X STOP procedure. Because the X STOP is not implanted adjacent to nerve roots or the spinal cord, the risk of neurologic deficit or paralysis may be considered minimal, and no incidence of either complication was reported in this study. Compared with the incidence and severity of complications cited in the laminectomy literature, the complications associated with the X STOP procedure suggest that the procedure is at least as safe as a decompressive laminectomy, and likely safer. In addition, the X STOP does not result in any significant radiographic changes to the lumbar spine. There were no differences between the mean disc height, curvature of the spine, or angulation of the spine of X STOP and control patients compared at 1 and 2 years. There was also no difference in the degree of spondylolisthesis between the X STOP and the control groups.

The incidence of a second operation for unresolved stenosis symptoms in the X STOP group was 6% through 2-year follow-up, a rate of reoperation favorably comparable with rates reported in the clinical literature for the surgical treatment of stenosis.<sup>29,30,32,54</sup> Atlas *et al*<sup>32</sup> reported a 6% reoperation rate at 1 year follow-up for 81 patients. Markwalder *et al*<sup>29</sup> reported a reoperation rate of 12% (12 of 100 patients) at a mean follow-up period of 2.9 years, and Jonsson *et al*<sup>30</sup> reported a reoperation rate of 18% (19 of 105 patients) occurring from 0.5 years to 4.5 years after the initial operation. Katz *et al*<sup>54</sup> reported a reoperation rate of 6%



at 1 year follow-up (5 of 88), which increased to 17% at a median follow-up period of 4.2 years.

Outcomes in the control group were significantly worse than those reported in the clinical literature for nonoperative therapy. However, the low success rate for nonoperative therapy should be considered a result of the rigorous outcomes measure used in the study, and not a confirmation that nonoperative therapy is not efficacious. NIC patients are typically considered as successes in the clinical literature if they experience at least some improvement after undergoing nonoperative therapy.<sup>10,26,32,45,46,55</sup> Hurri *et al*<sup>10</sup> reported that 44% had at least some improvement in neurologic symptoms, and Atlas *et al*<sup>32</sup> reported 45% had improvement in leg pain. Johnsson *et al*<sup>45</sup> found that 32% of the patients treated nonoperatively considered their condition improved. In this trial, 44% of control patients experienced at least some improvement in pain symptoms and 43% experienced some improvement in their physical function. The outcomes of the control group in this study were consistent with and comparable with results reported in the literature.

The genesis of the concept that an implant placed between the spinous processes might provide relief for patients suffering from neurogenic intermittent claudication came about from a straightforward clinical observation; most of these patients get relief of symptoms when they bend forward and flex their spines and conversely their symptoms worsen when they stand erect and extend their spines. Results of this randomized, multicenter trial clearly demonstrate that the X STOP improves clinical symptoms and function significantly compared with epidural steroid injections and conservative therapy in patients with symptoms of NIC. In each domain of the primary outcomes measure, X STOP patients had significantly better outcomes at every follow-up visit. The absence of any major complications demonstrates that the X STOP is safe. Because the X STOP procedure may be performed with a small exposure under local anesthesia, this treatment represents an attractive alternative for NIC patients.

The X STOP provides a conservative, yet effective, treatment for patients suffering from lumbar spinal stenosis. In the continuum of treatment options, the X STOP offers an attractive alternative to both nonoperative treatment and decompression surgery for patients with symptoms related to lumbar spinal stenosis.

### ■ Key Points

- A randomized, controlled, prospective multicenter trial of neurogenic intermittent claudication patients was conducted to compare the safety and efficacy of the X STOP interspinous implant with nonoperative therapy.

- Using a validated, patient-completed, condition-specific outcomes measure, the efficacy of the X STOP treatment was significantly greater than the control group, yet with a comparably low complication rate.
- The X STOP is a safe and effective treatment for neurogenic intermittent claudication patients compared with both nonoperative therapy and decompressive surgery.

### References

- Long DM, BenDebba M, Torgerson WS, et al. Persistent back pain and sciatica in the United States: patient characteristics. *J Spinal Disord* 1996;9:40–58.
- Fanuele JC, Birkmeyer NJ, Abdu WA, et al. The impact of spinal problems on the health status of patients: have we underestimated the effect? *Spine* 2000;25:1509–14.
- Boden SD, Davis DO, Dina TS, et al. Abnormal magnetic-resonance scans of the lumbar spine in asymptomatic subjects. A prospective investigation. *J Bone Joint Surg Am* 1990;72:403–8.
- Hart LG, Deyo RA, Cherkin DC. Physician office visits for low back pain. Frequency, clinical evaluation, and treatment patterns from a U.S. national survey. *Spine* 1995;20:11–9.
- Grazier KL, Holbrook TL, Kelsey JL, et al. *The Frequency of Occurrence, Impact, and Cost of Musculoskeletal Conditions in the United States*. Chicago, IL: American Academy of Orthopaedic Surgeons, 1984.
- Deyo RA. Promises and limitations of the Patient Outcome Research Teams: the low-back pain example. *Proc Assoc Am Physicians* 1995;107:324–8.
- Arbit E, Pannullo S. Lumbar stenosis: a clinical review. *Clin Orthop* 2001;384:137–43.
- Blau JN, Logue V. The natural history of intermittent claudication of the cauda equina. A long term follow-up study. *Brain* 1978;101:211–22.
- Dyck P. The stoop-test in lumbar entrapment radiculopathy. *Spine* 1979;4:89–92.
- Hurri H, Slati P, Soini J, et al. Lumbar spinal stenosis: assessment of long-term outcome 12 years after operative and conservative treatment. *J Spinal Disord* 1998;11:110–5.
- Iguchi T, Kurihara A, Nakayama J, et al. Minimum 10-year outcome of decompressive laminectomy for degenerative lumbar spinal stenosis. *Spine* 2000;25:1754–9.
- Inufusa A, An HS, Lim TH, et al. Anatomic changes of the spinal canal and intervertebral foramen associated with flexion-extension movement. *Spine* 1996;21:2412–20.
- Jenis LG, An HS. Spine update. Lumbar foraminal stenosis. *Spine* 2000;25:389–94.
- Johnsson KE, Rosen I, Uden A. The natural course of lumbar spinal stenosis. *Clin Orthop* 1992;279:82–6.
- Jonsson B, Annertz M, Sjöberg C, et al. A prospective and consecutive study of surgically treated lumbar spinal stenosis. Part I: clinical features related to radiographic findings. *Spine* 1997;22:2932–7.
- Schonstrom N, Willen J. Imaging lumbar spinal stenosis (review). *Radiol Clin North Am* 2001;39:31–53.
- Verbiest H. A radicular syndrome from developmental narrowing of the lumbar vertebral canal. *J Bone Joint Surg* 1954;36B:230–7.
- Verbiest H. Fallacies of the present definition, nomenclature, and classification of the stenoses of the lumbar vertebral canal. *Spine* 1976;1:217–25.
- Penning L, Wilmink JT. Posture-dependent bilateral compression of L4 or L5 nerve roots in facet hypertrophy. A dynamic CT-myelographic study. *Spine* 1987;12:488–500.
- Porter RW. Spinal stenosis and neurogenic claudication. *Spine* 1996;21:2046–52.
- Willen J, Danielson B, Gaulitz A, et al. Dynamic effects on the lumbar spinal canal: axially loaded CT-myelography and MRI in patients with sciatica and/or neurogenic claudication. *Spine* 1997;22:2968–76.
- Sortland O, Magnaes B, Hauge T. Functional myelography with metrizamide in the diagnosis of lumbar spinal stenosis. *Acta Radiol Suppl* 1977;355:42–54.
- Chung SS, Lee CS, Kim SH, et al. Effect of low back posture on the morphology of the spinal canal. *Skeletal Radiol* 2000;29:217–23.

24. Dong G, Porter RW. Walking and cycling tests in neurogenic and intermittent claudication. *Spine* 1989;14:965-9.
25. Schonstrom N, Lindahl S, Willen J, et al. Dynamic changes in the dimensions of the lumbar spinal canal: an experimental study in vitro. *J Orthop Res* 1989;7:115-21.
26. Amundsen T, Weber H, Lilleas F, et al. Lumbar spinal stenosis. Clinical and radiologic features. *Spine* 1995;20:1178-86.
27. Ciol MA, Deyo RA, Howell E, et al. An assessment of surgery for spinal stenosis: time trends, geographic variations, complications, and reoperations. *J Am Geriatr Soc* 1996;44:285-90.
28. Deyo RA, Cherkin DC, Loeser JD, et al. Morbidity and mortality in association with operations on the lumbar spine. The influence of age, diagnosis, and procedure. *J Bone Joint Surg Am* 1992;74:536-43.
29. Markwalder TM. Surgical management of neurogenic claudication in 100 patients with lumbar spinal stenosis due to degenerative spondylolisthesis. *Acta Neurochir (Wien)* 1993;120:136-42.
30. Jonsson B, Stromqvist B. Lumbar spine surgery in the elderly. Complications and surgical results. *Spine* 1994;19:1431-5.
31. Katz JN, Stucki G, Lipson SJ, et al. Predictors of surgical outcome in degenerative lumbar spinal stenosis. *Spine* 1999;24:2229-33.
32. Atlas SJ, Deyo RA, Keller RB, et al. The Maine Lumbar Spine Study, Part III. 1-year outcomes of surgical and nonsurgical management of lumbar spinal stenosis. *Spine* 1996;21:1787-94; discussion 94-5.
33. Turner JA, Ersek M, Herron L, et al. Surgery for lumbar spinal stenosis. Attempted meta-analysis of the literature. *Spine* 1992;17:1-8.
34. Zucherman JF, Hsu KY, Hartjen CA, et al. A prospective randomized multicenter study for the treatment of lumbar spinal stenosis with the X STOP interspinous implant: 1-year results. *Eur Spine J* 2004;13:22-31.
35. Lindsey DP, Swanson KE, Fuchs P, et al. The effects of an interspinous implant on the kinematics of the instrumented and adjacent levels in the lumbar spine. *Spine* 2003;28:2192-7.
36. Wiseman C, Lindsey DP, Fredrick AD, et al. The effect of an interspinous process implant on facet loading during extension. *Spine* 2005;15;30:903-7.
37. Swanson KE, Lindsey DP, Hsu KY, et al. The effects of an interspinous implant on intervertebral disc pressures. *Spine* 2003;28:26-32.
38. Richards JC, Majumdar S, Lindsey DP, et al. The treatment mechanism of an interspinous process implant for lumbar neurogenic intermittent claudication. *Spine* 2005;30:744-9.
39. Stucki G, Liang MH, Fossel AH, et al. Relative responsiveness of condition-specific and generic health status measures in degenerative lumbar spinal stenosis. *J Clin Epidemiol* 1995;48:1369-78.
40. Stucki G, Daltroy L, Liang MH, et al. Measurement properties of a self-administered outcome measure in lumbar spinal stenosis. *Spine* 1996;21:796-803.
41. Neumann P, Wang Y, Karrholm J, et al. Determination of inter-spinous process distance in the lumbar spine. Evaluation of reference population to facilitate detection of severe trauma. *Eur Spine J* 1999;8:272-8.
42. AHRQ. Treatment of degenerative lumbar spinal stenosis. Rockville, MD: Agency for Healthcare Research and Quality 2001:1-265.
43. Katz JN. Spinal Stenosis Data. Boston: Harvard Medical School, 2003:1-33.
44. Gunzburg R, Keller TS, Szpalski M, et al. Clinical and psychofunctional measures of conservative decompression surgery for lumbar spinal stenosis: a prospective cohort study. *Eur Spine J* 2003;12:197-204.
45. Johnsson KE, Uden A, Rosen I. The effect of decompression on the natural course of spinal stenosis. A comparison of surgically treated and untreated patients. *Spine* 1991;16:615-9.
46. Atlas SJ, Keller RB, Robson D, et al. Surgical and nonsurgical management of lumbar spinal stenosis: four-year outcomes from the Maine lumbar spine study. *Spine* 2000;25:556-62.
47. Benz RJ, Ibrahim ZG, Afshar P, et al. Predicting complications in elderly patients undergoing lumbar decompression. *Clin Orthop* 2001;384:116-21.
48. Khoo LT, Fessler RG. Microendoscopic decompressive laminotomy for the treatment of lumbar stenosis. *Neurosurgery* 2002;51:146-54.
49. Reindl R, Steffen T, Cohen L, et al. Elective lumbar spinal decompression in the elderly: is it a high-risk operation? *Can J Surg* 2003;46:43-6.
50. Postacchini F, Cinotti G, Perugia D, et al. The surgical treatment of central lumbar stenosis. Multiple laminotomy compared with total laminectomy. *J Bone Joint Surg Br* 1993;75:386-92.
51. Rosomoff HL. Neural arch resection for lumbar spinal stenosis. *Clin Orthop* 1981;83-9.
52. Fox MW, Onofrio BM, Hanssen AD. Clinical outcomes and radiological instability following decompressive lumbar laminectomy for degenerative spinal stenosis: a comparison of patients undergoing concomitant arthrodesis versus decompression alone. *J Neurosurg* 1996;85:793-802.
53. Kalbarczyk A, Lukes A, Seiler RW. Surgical treatment of lumbar spinal stenosis in the elderly. *Acta Neurochir (Wien)* 1998;140:637-41.
54. Katz JN, Lipson SJ, Chang LC, et al. Seven- to 10-year outcome of decompressive surgery for degenerative lumbar spinal stenosis. *Spine* 1996;21:92-8.
55. Simotas AC. Nonoperative treatment for lumbar spinal stenosis. *Clin Orthop* 2001;384:153-61.

# Configuration of the Connective Tissue in the Posterior Atlanto-Occipital Interspace

## A Sheet Plastination and Confocal Microscopy Study

Lance Nash, MSc,\* Helen Nicholson, MB, PhD,\* Antonio S. J. Lee, MSc,\*  
Gillian M. Johnson, PhD,† and Ming Zhang, MB, MMed, PhD\*

**Study Design.** The connective tissue structures in the posterior atlanto-occipital region were investigated using E12 sheet plastinations and confocal microscopy.

**Objectives.** To define the relationship between rectus capitis posterior minor (RCPm), posterior atlanto-occipital (PAO) membrane, nuchal ligament, and the spinal dura in the PAO interspace.

**Summary of Background Data.** It has been speculated that connections between the dura and muscles and/or ligaments in the PAO interspace may transmit forces from the cervical spine joint complexes to the pain-sensitive dura, generating cervicogenic headaches. Anatomic structures involved in these connections include the RCPm, PAO membrane, and nuchal ligament. However, there is little information about the nature of these connections and the relationships between these anatomic structures.

**Methods.** The study used a combined approach, consisting of the gross anatomic dissection of nine cadavers and the E12 sheet plastination method for thirteen adult human cadavers, five of which were further examined using confocal microscopy.

**Results.** The study demonstrates that (1) the tendinous fibers from the medial and deep part of the RCPm muscle are continuous antero-inferiorly with the spinal dura; (2) the PAO membrane is part of the RCPm fascia and tendon and the perivascular sheathes; (3) antero-inferiorly the PAO membrane fuses with the spinal dura rather than the atlas; and (4) the nuchal ligament does not exist in the PAO interspace.

**Conclusions.** The connective tissue structures that connect the spinal dura to the RCPm muscle in the PAO interspace are the RCPm fascia and tendinous fibers and perivascular sheathes.

**Key words:** fibrous connections, spinal dura, posterior atlanto-occipital membrane, rectus capitis posterior minor, posterior atlanto-occipital interspace, cervical spine.

**Spine 2005;30:1359–1366**

ists that dura-muscular and dura-ligamentous connections transmit forces from the cervical spine joint complexes to the pain-sensitive spinal dura and may contribute to the pathologic development of cervicogenic headaches (reviewed by Alix<sup>9</sup>). It is generally accepted that the RCPm muscle and the nuchal ligament are attached, via a connective tissue bridge, to the PAO membrane rather than directly to the spinal dura.<sup>1,7</sup> The PAO membrane, extending between the posterior arch of the atlas (C1) and the occiput, has been observed to tether the spinal dura by numerous connective tissue fibers creating a PAO membrane-spinal dura complex.<sup>1</sup> Several studies have reported that the connective tissue bridge originates from the connective tissue of the RCPm muscle.<sup>1,5</sup> More recently, this view has been further clarified by Dean and Mitchell<sup>7</sup> who found that a connective tissue bridge existed between the nuchal ligament and spinal dura rather than between the RCPm muscle and spinal dura. However, Johnson *et al*<sup>10</sup> and Zhang and Lee<sup>11</sup> have also reported that above the C2 level, the nuchal ligament can not be clearly identified, particularly in the deep region where the posterior midline is filled with fatty tissue and small vascular structures. Therefore, further investigations are needed to identify the origin of this connective tissue bridge and to clarify the relationship between the spinal dura, RCPm muscle, PAO membrane, and nuchal ligament in the PAO interspace.

The study of the connective tissue configuration in a cadaver is problematic and great difficulties exist in dissecting out the fasciae. Under a dissecting microscope, it is possible to trace the aponeurotic or tendinous fibers of a muscle. However, it is almost impossible to distinguish between the membranous (or fibrous) part of the subcutaneous tissue, deep fascia, epimysium, and epitenidium. Although the use of histologic methods to examine these fibrous tissues may overcome the problem, the application of such methods is greatly limited by the size of the sample areas. The E12 sheet plastination technique<sup>12</sup> provides a new approach to illustrate the detailed structural arrangement of connective tissue at both the macroscopic and microscopic levels. Furthermore, a recent study has demonstrated that the autofluorescence from collagen fibers in a plastinated specimen can be specifically detected by confocal microscopy.<sup>13</sup> Collagen fibers are one of the main components of connective tissue. Thus, the aim of this study was to use the sheet plastination technique in conjunction with confocal microscopy to reveal the relationship between the various

Recently, the spinal dural connection of the rectus capitis posterior minor (RCPm) muscle<sup>1–6</sup> and the nuchal ligament<sup>7,8</sup> in the posterior atlanto-occipital (PAO) interspace has been reported extensively. Speculation ex-

From the \*Department of Anatomy and Structural Biology and the †School of Physiotherapy, University of Otago, Dunedin, New Zealand.

Acknowledgment date: May 18, 2004. Acceptance date: July 28, 2004. The manuscript submitted does not contain information about medical device(s)/drug(s).

Institutional funds were received in support of this work. No benefits in any form have been or will be received from a commercial party related directly or indirectly to the subject of this manuscript.

Address correspondence and requests for reprints to Ming Zhang, Department of Anatomy & Structural Biology, University of Otago, PO Box 913, Dunedin, 9001 New Zealand; E-mail: ming.zhang@stonebow.otago.ac.nz



connective tissue structures in the PAO interspace which include the RCPm tendon and fascia, perivascular sheathes, PAO membrane, spinal dura, and periosteum.

## ■ Materials and Methods

Twenty-two human adult cadavers (16 males and 6 females, aged 65–89 years) were used in this study. The bodies were fixed with 10% neutral buffered formalin within 24 hours of death. They were bequeathed under the Human Tissue Act of 1964 to the Department of Anatomy and Structural Biology at the Otago School of Medical Sciences for research and teaching.

The head and neck from 13 cadavers (10 males and 3 females, aged 67–89 years) were processed as sheet plastinations (3 transverse, 2 coronal, and 8 sagittal) according to the E12 Sheet Plastination technique (van Hagen, 1979). This technique preserves and fixes the cellular constituents of tissue in situ by removing lipids and water and replacing them with a curable polymer resin.<sup>12</sup> The process involves setting the tissue in 20% gelatin at  $-80^{\circ}\text{C}$  for 24 hours. The frozen specimens were positioned in one of three planes of orientation (coronal, sagittal, or transverse) and cut using a bandsaw with a 1.6-mm blade set at a thickness of 2.5 mm. The slices were immersed in acetone (86.5–100%) at  $-25^{\circ}\text{C}$  over an 8-week period for dehydration. They were then degreased at room temperature ( $18\text{--}24^{\circ}\text{C}$ ) for an additional 4 weeks. The processed sections were polymerized under vacuum pressure with epoxy resin [Biodur, E12/E1/AE10/AE30 (100:28:20:5), Rathausstrasse 18, 69126, Heidelberg, Germany] at  $0^{\circ}\text{C}$  for 24 hours. The polymerized slices were laminated between two 50  $\mu\text{m}$ -thick plastic sheets. This was done to separate and protect the impregnated tissue and allowed curing for 1 week at room temperature or in an oven at  $30\text{--}40^{\circ}\text{C}$  for 3 days to accelerate the process. The translucent plastinations were examined under a Leica MZ8 Stereoscopic Dissecting Microscope at magnification ranging from 1.25 to 5 times. Photographs were captured with a Nikon Coolpix 990 Digital Camera.

The plastination process results in connective tissue, especially collagen, being endogenously autofluorescent at the 488-nm excitation<sup>13</sup>. Five sagittal slices from three cadavers were examined using a Biorad confocal laser-scanning microscope. The thickness of the optical section was set up at 107  $\mu\text{m}$  under a 5 times objective, and the images were electronically recorded.

Gross anatomic dissections were performed on 9 cadavers (6 males and 3 females, aged 65–89 years). Six of these cadavers (4 males and 2 females) were sectioned along the midsagittal plane from the occiput to C7 and one side of each cadaver was dissected according to a procedure described in previous studies.<sup>1,7,8</sup> The remaining 3 cadavers (2 males and 1 female) were examined by a progressive layer-by-layer dissection. Special attention was given to the deep posterior cervical region, RCPm muscle and its fascia, vertebral venous plexus, and PAO membrane. Gross anatomic images were recorded with a Nikon Coolpix 990 Digital Camera.

## ■ Results

In this study, the term fascia was defined as condensations of connective tissue on the surfaces of a muscle (epimysium) and its tendon (epitendinium). Compared with other connective tissue, fascial fibers are arranged more compactly and with a high degree of regularity in their directions (Figure 1a) and they are directly contin-

uous with the connective tissue within the muscle (perimysium; Figure 1a). As indicated in Figure 1a, the lateral and medial areas of the PAO interspace and the deep nuchal (posterior median) region were the focus of this study.

### ***The Tendinous Fibers of the RCPm Muscle Fuse Antero-Inferiorly With the Spinal Dura via the PAO Interspace***

Near the occiput, the RCPm tendinous fibers were extremely short, and the muscular fibers arose almost directly from the periosteum on the posterior aspect of the occiput (Figure 2a). Antero-inferiorly, the RCPm tendinous fibers were long and continuous with the periosteum of the posterior arch of the atlas (Figures 2a and b). The attachment of the RCPm tendon on the posterior arch appeared “L” shaped. Laterally the RCPm tendinous fibers were mainly attached to the posterior aspect of the arch (Figure 2b) whereas medially they fully covered the anterior, superior, and posterior surfaces of the arch (Figure 2a).

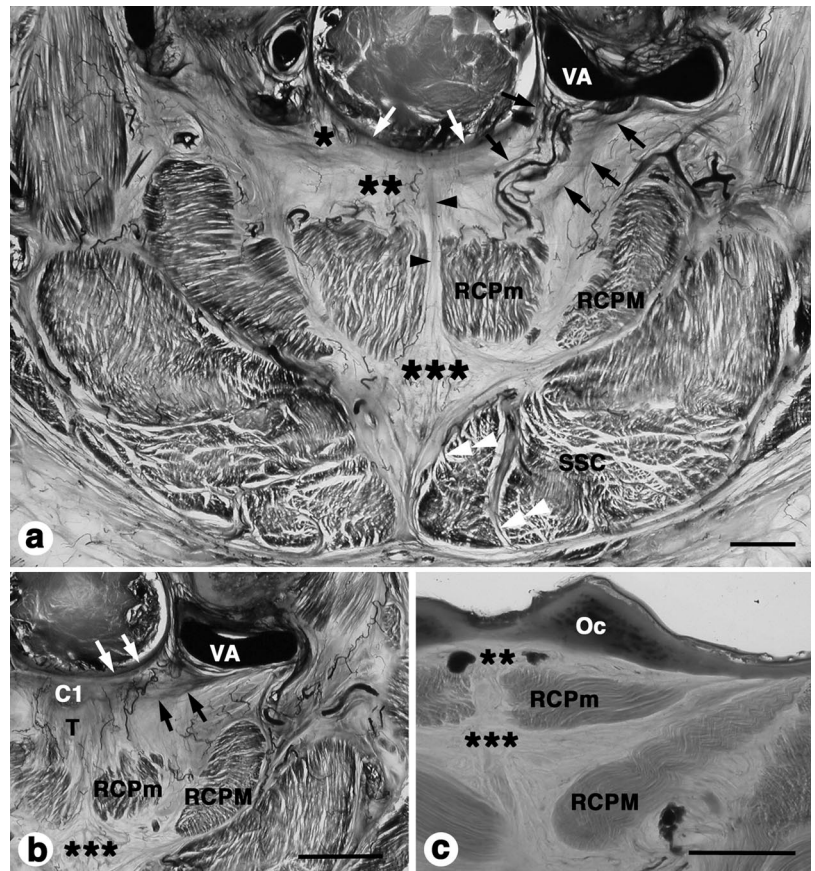
To examine the relationship between the RCPm muscle and the spinal dura in the PAO interspace, we specifically traced those RCPm tendinous fibers running toward the anterior aspect of the posterior arch of the atlas. In the sagittal slices, under a dissecting microscope, some of these RCPm tendinous fibers appeared to continue with the spinal dura just anterior to the posterior arch of the atlas (Figure 2a). Since the epidural space between the periosteum of the atlas and the spinal dura normally contains the internal vertebral venous plexus, the tendinous fibers identified previously may continue with the perivascular sheath of the plexus rather than the dura itself. To clarify this possibility, we examined a further five sagittal slices from three cadavers using confocal microscopy. Figure 3a clearly demonstrates that the RCPm tendinous fibers are directly continuous with the spinal dura via the PAO interspace and become a part of the spinal dura.

However, compared with the main part of the RCPm tendon, the tendinous fibers involved in the above RCPm tendon-spinal dural connection were very small and limited to those in the deep and medial part of the tendon (Figure 3a). At the gross anatomy level, it was almost impossible to distinguish these tendinous fibers from the RCPm fascia and the perivascular sheath of the external vertebral plexus (see below for further description).

### ***The PAO Membrane Is Formed by the RCPm Fascia and Vertebral Vascular Sheath and Antero-Inferiorly Fuses With the Spinal Dura***

In the transverse and sagittal slices from 11 cadavers, membrane-like structures were found in two regions of the PAO-interspace. Laterally, a double-layered membrane extended between the RCPm muscle and vertebral vascular sheath (Figures 1a and 2d). Under the confocal microscope, this membrane was found to continue directly with the muscular fibers of the RCPm (Figure 3c).

Figure 1. Horizontal (**a,b**) and coronal (**c**) views of the posterior atlanto-occipital (PAO) interspace in the plastinated slices. The lateral and medial areas of the PAO interspace and the pyramidal fatty space in the deep nuchal region were the three main areas examined in this study and are marked with single, double, and triple asterisks, respectively. Slice (**a**) is through the upper PAO interspace where the vertebral artery (VA) lies in contact with the dural mater (white arrows). Black arrows point to the lateral part of the PAO membrane that is double-layered and directly continuous with the perivascular sheath of the VA. A small vascular bundle can be seen between the double layers of the PAO membrane and supplies the rectus capitis posterior minor (RCPm). The black arrowheads point to the medial part of the RCPm fascia that runs anteriorly toward the spinal dura and postero-laterally continues with the fascia of the rectus capitis posterior major (RCPM). The white double arrowheads in the right semi-spinalis capitis (SSC) indicate the continuity between the epimysium (the fascia) and the perimysium within the muscle. Slice (**b**) is through the PAO interspace, just above the upper margin of the atlas (C1). The attachment of the RCPm tendon to the atlas is marked with "T." The black and white arrows point to the PAO membrane and spinal dura, respectively. Slice (**c**) is a coronal section just posterior to the posterior margin of the foramen magnum, showing the continuity of fatty tissue between the medial part of the PAO interspace (double asterisks) and the pyramidal fatty space (triple asterisks). Oc, occiput. Scale bar = 5 mm.



The membrane thus anchored the RCPm muscle to the vertebral vascular sheath (Figure 2d) and was then, via the vascular sheath, continuous with the spinal dura (Figure 1a and b).

Medially, deep to the RCPm muscle, the double-layered membrane could not be clearly identified (Figures 1a and b, 2a and b). Following the serial sagittal sections from the medial to the lateral of the PAO interspace (Figure 2), the medial part of the double-layered membrane appeared to split in order to accommodate the external vertebral plexus and its surrounding connective tissue (Figures 2a and b). The deep layer of the membrane fused inferiorly with the spinal dura whereas the superior layer continued with the deep layer of the RCPm fascia (Figures 2a and b). However, both layers were very thin and were often interrupted by the perivascular sheaths of the external vertebral plexus (Figures 1a and b, 2a and b). This made it difficult to distinguish them from each other (Figure 2b), particularly with regard to the RCPm fascia because some of these vascular structures ran into the deep aspect of the RCPm muscle and supplied the muscle (Figure 1a). Inferiorly, the RCPm fascia and the perivascular sheath were also fused with the spinal dura (Figures 2a, 3a and b). This fascial-

dura connection was particularly apparent medially (Figures 1a, 2a, 3a and b).

In addition to the external vertebral plexus, other vascular structures in the PAO-interspace were the internal vertebral plexuses and the marginal dural sinus. Based on their relationship with the dura, it was possible to localize them separately on the sagittal sections. Under the confocal microscope, the spinal dura in the PAO interspace appeared "V" shaped and consisted of three layers. These layers arose from the periosteum of the posterior, inferior, and anterior aspects of the occiput (Figure 3b). The internal vertebral venous plexus was sandwiched between these three layers (Figures 2b and 3b) and superiorly communicated with the marginal dural sinus (Figure 3b). The external vertebral plexus was posterior to the dura (Figures 2a,b and 3b).

In summary, the so-called "PAO membrane" was formed by the RCPm tendon (or aponeurosis) laterally (Figures 2d and 3c) and the RCPm fascia and the perivascular sheaths of vertebral vessels medially (Figures 2a and b). Superiorly, the membrane continued with the periosteum of the occiput and inferiorly fused with the spinal dura and the perivascular sheath of the internal vertebral plexus. Therefore, the lateral part of the PAO



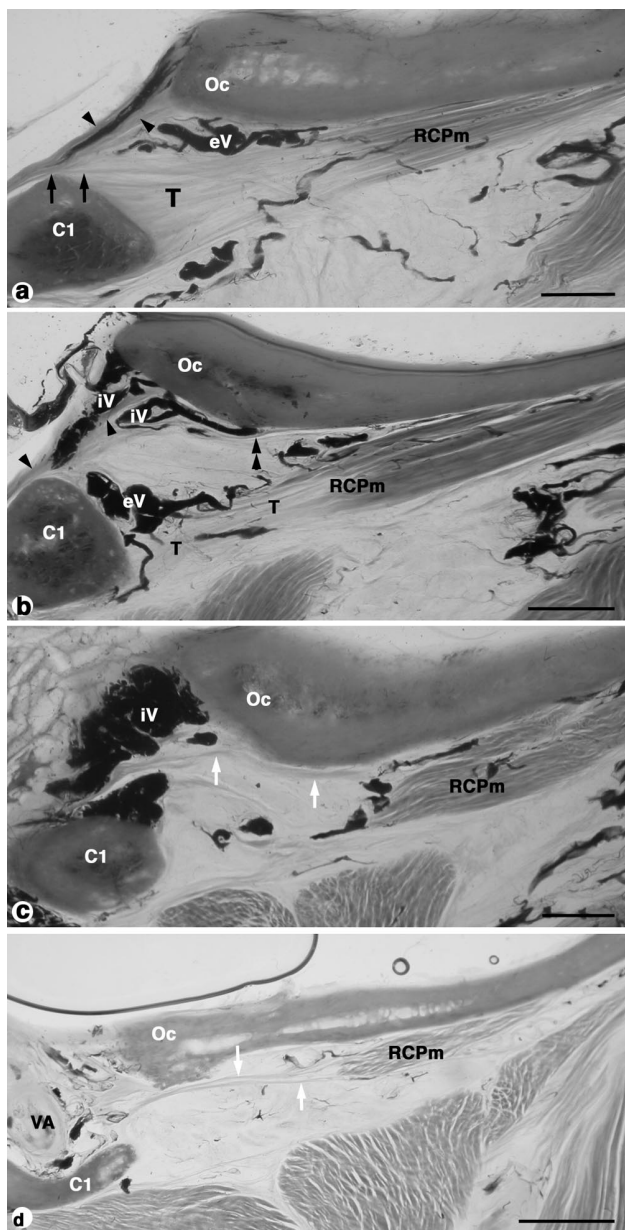


Figure 2. A series of sagittal plastinated slices from the posterior atlanto-occipital (PAO) region commencing from the midline (a) to the lateral aspect (d). The black arrows point to small bundles of tendinous fibers (T) of the rectus capitis posterior minor (RCPm) that run towards the spinal dura (black arrowheads). The external vertebral venous plexus (eV) is located between the RCPm muscle and the posterior layer of the spinal dura [double arrowheads in (b)], whereas the internal vertebral plexus (iv) is either within the multiple layered spinal dura [black arrowheads and double arrowheads in (a,b)] or deep to the spinal dura (c). The white arrows label the lateral part of the PAO membrane that is double layered (d) and directly continues with the RCPm muscular fibers (d; see Figure 3 for the further evidence from confocal microscopy) and the perivascular sheath of the vertebral artery (VA; d). Oc, occiput; C1, atlas. Scale bar = 5 mm.

membrane appeared more obvious than the medial part and was relatively easy to be demonstrated at gross anatomy level (Figures 4b and c). The medial part of the membrane contained numerous venous plexuses and

fatty tissues and did not appear as a membrane-like structure (Figures 2a and b).

### ***The Nuchal Ligament Does Not Exist in the PAO Interspace***

No direct continuity between the nuchal ligament and the posterior cervical dura mater at the PAO interspace was found in this study. Along the posterior median plan, the superficial and deep muscular structures at this level were fully separated by a pyramidal space that was filled with fatty tissue containing small vascular structures (Figures 1a-c and 4a). This finding is in agreement with the previous reports.<sup>10,11</sup> The lateral extension of this space also separated the superficial layer of the RCPm fascia from the deep layer of RCPm fascia (Figures 1a and b). Figure 4a shows this fatty space at the gross anatomic level. There was no obvious perpendicular connective tissue sheet along the posterior midline region. The lateral extension of the fatty space between the deep layer of the RCPm fascia and superficial layer of the RCPm fascia was identified at gross anatomic level as well (Figure 4a). Dissecting one side of the head and neck, we found that it was difficult to fully expose the fatty space. The medial parts of the RCPm and RCPm fascia (Figure 1a, 4a and c) could be easily misidentified as the nuchal ligament.

### **Discussion**

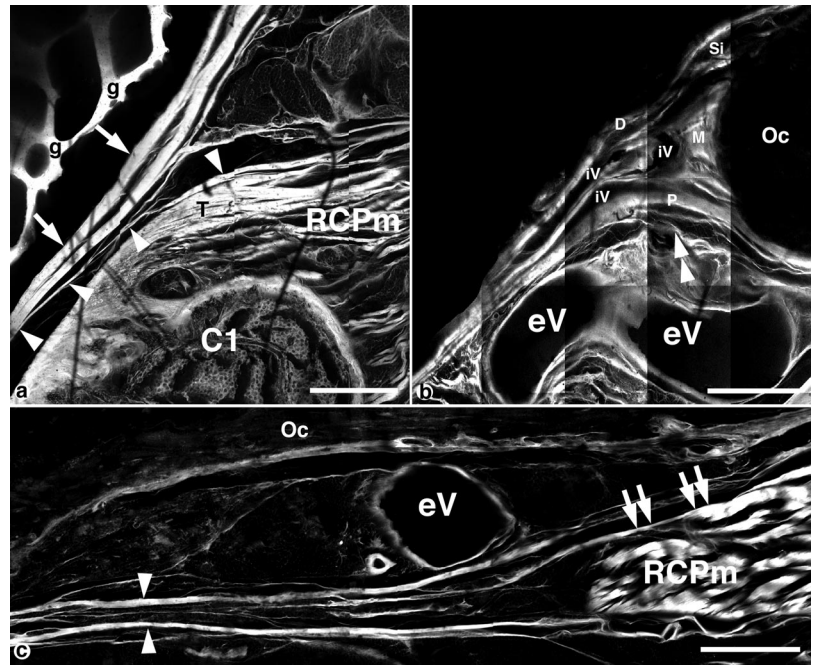
The PAO interspace is a very intricate area. The interspace contains the PAO membrane, transition zone between the cranial and spinal dura, perivascular sheaths associated with the internal and external vertebral venous plexuses and marginal dural sinus, and the deep layer of the RCPm fascia. Using a sheet plastination technique in conjunction with confocal microscopy, this study describes the detailed relationship between these connective tissue structures and clarifies the origins of the connective tissue bridge between the RCPm and spinal dura mater.

### ***The Tendon of the RCPm Directly Attaches to the Spinal Dura via the PAO Interspace***

In 1995, Hack et al<sup>1</sup> reported that a dense fibrous connective tissue bridge attached the deep surface of the RCPm muscle to the PAO membrane-spinal dural complex. Their finding was supported by several other studies,<sup>1-3,5</sup> including Dean and Mitchell<sup>7</sup> who also described a connective tissue connection between the nuchal ligament and spinal dura at the PAO interspace (see below for further discussion). However, the origin of this connective tissue bridge has not been clearly identified. There were no reports showing a direct continuity of individual connective tissue fibers or bundles from the RCPm muscle to the dura. Normally, there are three types of connective tissue associated with a skeletal muscle: tendon, epimysium, and epitendinium. Tendinous fibers directly continue with muscular fibers via myotendon junctions, whereas the epimysium and epitendinium envelop a muscle and continue with the perimy-



Figure 3. Confocal images of the lower (a) and upper (b) parts of the medial PAO interspace and lateral PAO membrane (c) in sagittal plastinated slices. Confocal microscopy was used to detect the autofluorescences from collagen fibers. The thickness of these optical sections were 107  $\mu\text{m}$ . Optical section (a) shows a small bundle of the tendinous fibers (T) of the rectus capitis posterior minor muscle (RCPm) fusing with the spinal dura (arrows). Optical section (b) demonstrates that the posterior uppermost part of the spinal dura is a multiple layered structure and can be divided into the posterior (P), middle (M), and anterior (D) layers that are continuous with the periosteum of the posterior, inferior, and anterior surfaces of the occiput (Oc), respectively. Among these layers are small vessels from the internal vertebral plexus (iV) that communicate with the dural sinus (Si) between the inner and external layers of the cranial dura. Note that the perivascular connective tissue (the white double arrowheads) of the external vertebral plexus (eV) serves as the medial PAO membrane and fuses with the posterior layer (P) of the spinal dura. Optical section (c) shows the double-layered lateral PAO membrane (arrowheads). The double arrows point to the junction between the muscular and tendinous portions of the RCPm muscle, indicating the part of the PAO membrane is the tendon of the RCPm muscle. C1, atlas; g, gelatin used for embedding. Scale bar = 5 mm.



sium and peritendinium inside the muscle and its tendons, respectively. The present study demonstrates, for the first time, that tendinous fibers of the RCPm muscle are continuous with the spinal dura (Figure 3a). These tendinous fibers mainly originate from the deepest part of the RCPm muscle (Figures 2a and 3a) whereas the majority of the tendinous fibers of the muscle inferiorly attach to the superior and posterior aspects of the posterior arch of the atlas (Figures 2a and b). A direct continuity of the RCPm muscle to the spinal dura should have a great impact on strengthening the dura and preventing dural enfolding.<sup>1</sup> Such an impact becomes particularly important during extension of the head and neck because, when the RCPm muscle extends the cranio-cervical junction, a small portion of its muscular fibers simultaneously contract to pull the spinal dura posteriorly, preventing dural enfolding. However, the microstrain and trauma in the RCPm muscle and its tendon may cause a clonous condition of the muscle and stimulate the pain-sensitive dura, generating a cervicogenic headache.

#### What Is the PAO Membrane?

It is generally regarded that the PAO membrane is a thin, broad membrane that extends between the posterior margin of the foramen magnum and the upper border of the posterior arch of the atlas, with free lateral borders arching over the vertebral arteries.<sup>14</sup> Thus, the membrane was believed to be continuous with the periosteum of the occiput and the atlas and to be sandwiched between the spinal dura and the RCPm muscle.<sup>15</sup> Our re-

sults suggest that this concept is not correct. The so-called “PAO membrane” is mainly formed by the RCPm tendinous fibers, the deep layer of the RCPm fascia, and perivascular connective tissue sheathes (Figures 2 and 3c). The membrane can be divided into two portions. Laterally, the membrane contains the tendinous fibers of the most lateral and deep part of the RCPm muscle (Figure 3c). These tendinous fibers are attached to the perivascular sheath of the vertebral artery and the vertebral cavernous sinus, which has a direct connection with the periosteum of the lateral part of the posterior arch of the atlas and spinal dura (Figures 1a and b, and 2d). Thus, this portion of the membrane appears as a strong membrane-like structure, particularly along its lateral margin (Figures 4b and c). Medially, such a membrane-like structure becomes less obvious mainly because of the increase of vascular structures and their associated perivascular connective tissue. Antero-inferiorly, the deep layer of the RCPm fascia and the vascular sheath were fused to the spinal dura, rather than attached to the posterior arch of the atlas. In summary, therefore, we believe that the RCPm tendon (deep and lateral portions only) and fascia were historically misrepresented by the so-called PAO membrane. Unlike at the anterior aspect of the cranio-cervical junction, no real ligamentous structure appears to exist at the posterior cranio-cervical junction. Thus, we strongly question whether the PAO membrane is an appropriate anatomic term to represent those connective tissue structures in the PAO interspace, particularly in the medial area.

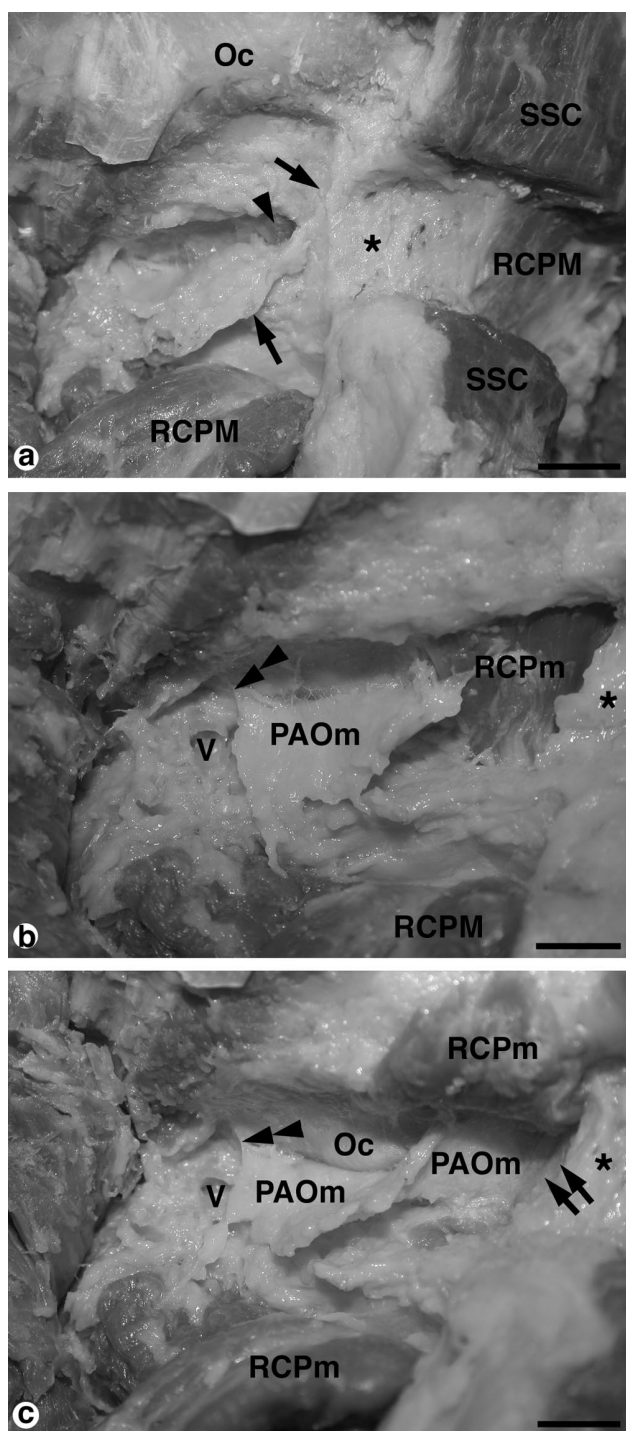


Figure 4. Gross anatomic views of the posterior atlanto-occipital (PAO) region. **a**, the fatty suboccipital space (asterisk, also in **b** and **c**) after a segmental removal of the semispinalis capitis (SSC). Note that no ligamentous structure can be identified in the fatty space. The left rectus capitis posterior major (RCPm) muscle has been reflected inferiorly to reveal the lateral extension of the fatty space between the RCPm fascia (arrows) and the superior surface of the rectus capitis posterior minor (RCPm; arrowhead). **b**, the lateral part of the PAO membrane (PAOm) after inferiorly reflecting the RCPm and removing the lateral extension of the fatty space. The double arrowhead points to the lateral border of the PAO membrane arching over a vertebral vessel (V; also in **b**). **c**, the medial part of the PAO membrane (PAOm) after reflecting the RCPm superiorly and inferiorly. The double arrows point to the medial part of the RCPm fascia that can be easily misidentified as part of the nuchal ligament. Oc, occiput. Scale bar = 5 mm.

The findings of this study contradict the general belief that the PAO membrane is the superior extension of the ligamentum flavum.<sup>14</sup> However, they strongly support several recent studies showing that the PAO membrane plays an insignificant role in cranio-cervical stability,<sup>16,17</sup> because inferiorly the membrane is not continuous with the periosteum of the atlas. Our results suggest that as the RCPm fascia is the main connective tissue structure in the PAO interspace, the RCPm fascia together with the muscle may make some contribution to the posterior cranio-cervical stability. Experimentally it may be a challenge to test the above assumption because a removal or incision of the RCPm muscle will destroy the integrity of the RCPm muscle and its fascia, and any further cut of the RCPm fascia (or the PAO membrane) will result in very little effect on the maintenance of stability of the specimen.

#### ***Is the Nuchal Ligament Attached to the PAO Membrane and Spinal Dura?***

Recently, after gross dissection of 10 adult cadavers, Dean and Mitchell<sup>7</sup> described a direct continuity between the nuchal ligament and posterior cervical dura at the atlanto-occipital junction. However, Johnson *et al*<sup>10</sup> reported regional differences of the connective tissue configuration along the nuchal ligament and claimed that the nuchal ligament could not be clearly distinguished from the surrounding connective tissue in the suboccipital region. Zhang and Lee<sup>11</sup> further demonstrated that a fatty potential space exists between the semispinalis fascia and the RCPm and RCPm fascia. In the present study, both the gross anatomy dissection and examination of sheet plastination further reveal that no obvious perpendicular fibers or bundles from the nuchal ligament traverse through the fatty suboccipital space (Figures 1a and 4a). As shown in Figure 1a, three connective tissue layers exist along the posterior midline in the suboccipital region: the semispinalis fascia, a fatty space, and the RCPm and RCPm fascia. The semispinalis fascia and RCPm fascia does contain some perpendicular fibers (Figure 1a), but they are not continuous with each other and are interrupted by the fatty space. This may be one reason why during gross dissection, the semispinalis fascia and the RCPm and RCPm fascia could be easily misidentified as a single ligamentous structure, particularly when the “nuchal ligament” was exposed by reflecting the posterior cervical muscles medially.

The RCPm fascia has a strong direct attachment to the periosteum of the external occipital crest and the posterior spinal dura along the posterior midline (Figures 1a and 4c). This medial portion of the RCPm fascia continues laterally with the deep layer of the RCPm fascia, which fuses with the perivascular sheath of the external vertebral plexus forming the medial part of the PAO membrane. Therefore, the fibrous connections between the nuchal ligament or PAO membrane to the spinal dura identified in sagittally dissected cadavers and magnetic resonance images<sup>1-8</sup> are most likely to be the RCPm



fascia. Surgically, the medial portion of the RCPm fascia may be a good candidate for duraplasty<sup>15,18</sup> because it is much stronger than other connective tissue structures in the region and is directly connected with the periosteum, which can be harvested together. If a large graft is needed, the superficial layer of the RCPm fascia, together with the RCPm fascia, can be included. On the basis of observations in this study, the deep layer of the RCPm fascia is not recommended for use in a dural graft. The main reason is that it merges with the external vertebral venous plexus.

### **The Multiple-Layered Cerebrospinal Dural Junction**

Historically, the spinal dura mater is described as a continuity of the inner layer of the cerebral dura mater and is attached to the circumference of the foramen magnum. There are few reports that describe the architecture of the cerebrospinal dural junction. In this study, we demonstrated that the upper-most part of the spinal dura at the PAO interspace arose from three sources: the periosteum of the anterior, inferior and posterior aspects of the occiput, and the inner layer of the cerebral dura mater. These four layers were normally separated from each other by small vascular structures. The upper spinal dura also received some small contributors, such as the RCPm tendinous fibers, RCPm fascia, and perivascular sheathes (Figures 3a and b). The multiple layered spinal dura mater (8 to 12 layers) has been reported before.<sup>19</sup> Interestingly, the fibers in the outermost layer of the spinal dura run in all three directions (longitudinal, horizontal and transverse), whereas the fibers in the other layers have a longitudinal orientation only.<sup>19</sup> The reasons for the multiple layered arrangement of the spinal dura is unclear. Our study suggests that in the PAO interspace, the orientation of fibers in the spinal dura may be related to the various sources of the dural layers.

Inferiorly, the multiple layers of the spinal dura in the PAO interspace gradually fused with each other and appeared as a “V” shape. Whether this is the reason for the posterior spinal dura being much thicker than the anterior dura in the upper cervical spine<sup>20</sup> is not clear. The thicker dura normally extends down to C3 level,<sup>20</sup> but there are no reports on the morphologic comparison between the anterior and posterior cerebro-spinal dura junction. The tent-like spinal dura in the PAO interspace was observed in magnetic resonance imaging but was presumed to be caused by previous hyperflexion injury.<sup>21</sup> The evidence from the present study strongly suggests that such interpretation needs to be revisited.

In conclusion, the deep layer of the RCPm fascia and the perivascular sheath of the external vertebral plexuses are the main connective tissue components between the RCPm muscle and spinal dura in the posterior cranio-cervical region. The RCPm fascia, perivascular sheathes, and a small portion of the RCPm tendinous fibers are attached antero-inferiorly to the spinal dura via the PAO interspace. The morphologic features of the RCPm tendon and fascia indicate that they may have an important

role in the maintenance of the posterior cranio-cervical stability and the prevention of the dural enfolding and are of anatomic relevance in the debate regarding the etiology of cervicogenic headaches.

### **■ Key Points**

- The RCPm tendon fuses with the spinal dura.
- The RCPm tendinous fibers and fascia and the perivascular sheathes form the PAO membrane.
- The PAO membrane fuses with the spinal dura.
- The nuchal ligament does not attach to the spinal dura.

### **Acknowledgments**

The authors would like to thank Andrew McNaughton, Otago Centre for Confocal Microscopy, University of Otago, for the confocal laser scanning. They would also like to thank Russell Barnett and Gavin Burland for technical assistance in the preparation of the E12 sheet plastinations.

### **References**

1. Hack GD, Koritzer RT, Robinson WL, et al. Anatomic relation between the rectus capitis posterior minor muscle and the dura mater. *Spine* 1995;20:2484–6.
2. Hack GD, Koritzer RT, Robinson WL, et al. The ligamentum craniale durae matris spinalis ligament may contribute to the prevention of dural enfolding [In response]. *Spine* 1997;22:925–6.
3. Humphreys BK, Kenin S, Hubbard BB, et al. Investigation of connective tissue attachments to the cervical spinal dura mater. *Clin Anat* 2003;16:152–9.
4. Lang J. The craniocervical junction - anatomy. In: Voth D, Glees P, editors. *Diseases in the Craniocervical Junction*. Berlin, NY: Walter de Gruyter, 1987:27–61.
5. Rutten HP, Szpak K, van Mameren H, et al. Anatomic relation between the rectus capitis posterior minor muscle and the dura mater. *Spine* 1997;22:924–6.
6. von Lanz T. Über die Rückenmarkshäute. 1. Die konstruktive Form der harten Haut des menschlichen Rückenmarkes und ihrer Bänder. *Arch Entwickl Mech Org* 1929;118:252–307.
7. Dean NA, Mitchell BS. Anatomic relation between the nuchal ligament (ligamentum nuchae) and the spinal dura mater in the craniocervical region. *Clin Anat* 2002;15:182–5.
8. Mitchell BS, Humphreys BK, O'Sullivan E. Attachments of the ligamentum nuchae to cervical posterior spinal dura and the lateral part of the occipital bone. *J Manipulative Physiol Ther* 1998;21:145–8.
9. Alix ME, Bates DK. A proposed aetiology of cervicogenic headache: the neurophysiologic basis and anatomic relationship between the dura mater and the rectus posterior capitis minor muscle. *J Manipulative Physiol Ther* 1999;22:534–9.
10. Johnson GM, Zhang M, Jones DG. The fine connective tissue architecture of the human ligamentum nuchae. *Spine* 2000;25:5–9.
11. Zhang M, Lee AS. The investing layer of the deep cervical fascia does not exist between the sternocleidomastoid and trapezius muscles. *Otolaryngol Head Neck Surg* 2002;127:452–4.
12. von Hagens G. Impregnation of soft biological specimens with thermosetting resins and elastomers. *Anat Rec* 1979;194:247–55.
13. Phillips MN, Nash LG, Barnett R, et al. The use of confocal microscopy for the examination of E12 sheet plastinated human tissue. *J Int Soc Plastination* 2003;17:12–6.
14. Yong-Hing K, Reilly J, Kirkaldy-Willis WH. The ligamentum flavum. *Spine* 1976;1:226–34.
15. Tubbs RS, Wellons JC, Blount JP, et al. Posterior atlantooccipital membrane for duraplasty. *J Neurosurg* 2002;97(2 Suppl):266–8.



16. Harris MB, Duval MJ, Davis JA Jr, et al. Anatomical and roentgenographic features of atlantooccipital instability. *J Spinal Disord* 1993;6:5–10.
17. Sun PP, Poffenbarger GJ, Durham S, et al. Spectrum of occipitoatlantoaxial injury in young children. *J Neurosurg* 2000;93(1 Suppl):28–39.
18. Kosnik EJ. Use of ligamentum nuchae graft for dural closure in posterior fossa surgery. *J Neurosurg* 1998;89:155–6.
19. Dittman M, Reina MA, Lopez Garcia A. New results in the visualization of the spinal dura mater with scanning electron microscopy. *Anaesthesist* 1998;47:409–13.
20. Taylor JR, Taylor MM, Twomey LT. Posterior cervical dura is much thicker than the anterior cervical dura [comment]. *Spine* 1996;21:2300–1.
21. Krakenes J, Kaale BR, Rorvik J, et al. MRI assessment of normal ligamentous structures in the craniovertebral junction. *Neuroradiology* 2001;43:1089–97.

# Recording of Neural Activity From Goat Cervical Facet Joint Capsule Using Custom-Designed Miniature Electrodes

Chaoyang Chen, MD, Ying Lu, MS, John M. Cavanaugh, MD, Srinivasu Kallakuri, MS, and  
Ajit Patwardhan, MD

**Study Design.** To establish a methodology for the neurophysiologic study of mechanoreceptors in the cervical facet joint capsule.

**Objectives.** To test a custom designed miniature dual bipolar electrode for recording the neural activity in cervical dorsal roots. To determine if the neural activity from different receptors in the capsule can be differentiated using this methodology.

**Summary of Background Data.** Injury to cervical facet joint capsules has been regarded as an important source of whiplash pain, but no neurophysiologic study has been performed to demonstrate or characterize sensory nerve function in the capsule.

**Methods.** Nineteen goats weighing 34 to 55 kg were used under general anesthesia. A C4–C6 laminectomy was performed to expose the C6 nerve root. Custom designed miniature dual bipolar electrodes were used to record neural activity in the left C6 branches. Electrical and mechanical stimuli were used to evoke receptor activity in the dorsal aspect of the C5/6 capsule. Conduction velocities (CVs) of evoked units were determined by electrical stimulation and dual-bipolar-electrode recording methods. The units were classified based on their CVs. The waveform of each classified unit was saved as a template for later single unit discharge search among multiunit discharges during the stretch of the capsule. The C5/6 facet joint with capsule was pulled by a computer-controlled actuator instrumented with a load cell at a rate of 0.5 mm per second. The evoked neural activity and load were recorded, digitized, and analyzed to determine CV, discharge rate, and response to the stretch.

**Results.** Miniature bipolar electrodes recorded the neural activity in both channels, with single unit CVs being measured. There was no discernible motion between the electrode and dorsal root when the capsule was pulled. Both local compression and stretch on capsule evoked multiunit discharges. A- $\beta$ , A- $\delta$ , and C-fiber units were found among these multiunit discharges. The rate of single unit and multiunit discharges increased during capsule stretch in the physiologic range and afterdischarges occurred beyond the physiologic range.

**Conclusions.** The novel miniature electrodes not requiring a micromanipulator made it feasible and reliable to record neural activity from short cervical spinal roots. Waveforms of different units could be identified, making it possible to study sensory functions of the facet joint capsule. A- $\beta$ , A- $\delta$ , and C-fiber units were found responding to mechanical stimuli, indicating that facet joint capsule has functional proprioceptors and nociceptors.

**Key words:** whiplash, cervical facet joint capsule, neurophysiology, nerve root, electrode. **Spine** 2005;30:1367–1372

Whiplash associated disorders are among the most common injuries associated with rear-end motor vehicle collisions. The estimated incidence of whiplash in the United States is 4 per 1000 people.<sup>1</sup> In 1991, the average cost of an Abbreviated Injury Scale 1 level neck injury was \$4000, with an annual cost of \$3.9 billion in the United States.<sup>2</sup> Approximately 25% of the patients experience persisting pain with 10% suffering serious pain.<sup>3</sup> The mechanisms of whiplash associated disorders are not clear. Recent clinical and biomechanical studies implicate the facet joints as perhaps the major source of persisting pain after whiplash.<sup>2,4–8</sup> The presence of nerve fibers<sup>9</sup> and nerve endings<sup>10</sup> in cervical facet joint capsules has also been reported. However, no neurophysiologic studies have been performed on cervical facet joint capsules to understand their neural response to mechanical stimulation.

Sheep have been used in biomechanical studies of cervical spine because of their comparability to the human cervical spine.<sup>8,11–13</sup> Goats have been used to assess the biomechanical strength of the developing cervical spine.<sup>14</sup> Goats were used in our current study to characterize neurophysiological response from cervical facet capsules. There were several challenges in undertaking this study: understanding the anatomy of the cervical spinal cord, adjacent nerve roots, and dorsal rami in the goat; placing a recording electrode under the relatively short cervical dorsal roots to obtain stable recordings; and applying tensile loads to the C5/6 facet joint capsule while maintaining a stable recording from the nerve roots. If successful this would be the first published study to record baseline neural activity from cervical facet capsules as well as under tensile loading conditions.

The principal purpose of this initial study was to demonstrate if a custom-designed miniature electrode could be used to record neural activity in the goat cervical dor-

From the Spine Research Laboratory, Bioengineering Center, Wayne State University, Detroit, Michigan.

Supported by CDC Grants R49 CCR519751 and R49-CE000455.

Acknowledgment date: December 1, 2003. First revision date: March 23, 2004. Acceptance date: August 2, 2004.

The manuscript submitted does not contain information about medical device(s)/drug(s).

Federal funds were received in support of this work. No benefits in any form have been or will be received from a commercial party related directly or indirectly to the subject of this manuscript.

Address correspondence and request for reprints to Chaoyang Chen, MD, Bioengineering Center, Wayne State University, 818 West Hancock Avenue, Detroit, MI 48202; E-mail: cchen@rrb.eng.wayne.edu

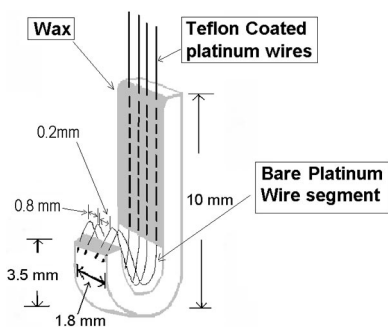


Figure 1. Design of miniature dual bipolar recording electrodes. Bare segments were protected by the wax hook from touching soft tissue other than the nerve root. The four ends of the bare wire segments were fixed in the wax so that they were kept stable inside the hook.

sal roots before and during stretch of the facet joint capsule, and if functional units from the identified receptive fields of the capsule could be characterized based on their conduction velocity and response to mechanical loading.

## Materials and Methods

**Preparation of Miniature Dual Bipolar Recording Electrode.** Four Teflon coated platinum wires (A-M Systems Inc., Carlsborg, WA) were used to fabricate dual bipolar electrodes. The coated wire had a diameter of 0.1 mm, and the wire with coating removed had a diameter of 0.05 mm. A 4 mm-long segment of each Teflon coat at the end was stripped off to expose the bare wire. This segment of wire was bent into "U" shape so as to be placed into the groove of a "J" shape wax mold. This wax mold was made of Jeweler's wax (Kindt-Collins Co., Cleveland, OH). The ends of bare segment were soldered into the wax, which held the wires stably so that the distances between electrodes remained constant. Two pairs of electrodes were installed into a wax mold. The distance between two wires was 0.8 mm. The distance between two pairs of electrodes was 0.2 mm. The wax bed was 1-mm thick and 2-mm wide. The long arm of the hook was 10 mm in length. The short arm of the hook was 3.5-mm long. Only the segments inside the hook were bare wires and in contact with nerve roots (Figure 1). The electrode ensheathed in wax was insulated from touching the surrounding soft tissue including the dural sac and spinal cord.

**Preparation of Animals.** Nineteen skeletally mature adult goats (17 female and 2 male) weighing 34 to 55 kg were used. Five goats were used to develop recording techniques. Fourteen goats were used for data acquisition. They were initially anesthetized by a combination of ketamine (4 mg/kg, intramuscular), diazepam (0.5 mg/kg, intramuscular), butorphenol (0.22 mg/kg, intramuscular) and atropine (0.066 mg/kg, intramuscular). Subsequently the anesthesia was maintained by the inhalation of isoflurane (2.5%). A C2 to T2 midline incision was made and splenius capitis, semispinalis capitis, and other multifidus muscles were separated from spinous processes and laminae. The ventral rami of C4–C5 were cut close to the end of the transverse process to block afferent neural activity from peripheral tissues. A C4–C6 laminectomy was then performed to expose the dural sac.

The dural sac was cut using a no. 11 blade and carefully pulled open to expose the C6 nerve branches. Two no. 4–

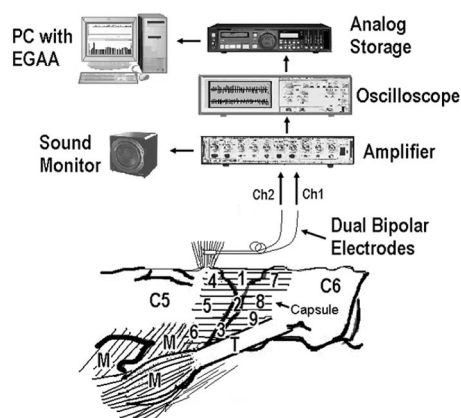


Figure 2. Experimental setting. C6 dorsal root branches are distributed in a fan shape between the foramen and spinal cord. The third branch of C6 dorsal root was usually picked up for recording. The capsule was divided into 9 areas and mechanically probed or electrically stimulated to evoke activity from receptors in the capsule. Neural activity was recorded in two channels. M, muscle; T, tendon; Ch1, Channel 1; Ch2, Channel 2.

sutures were anchored to the dural sac so as to pull it away to expose the nerve root branches. Using Dumont no. 5 forceps to hold the miniature electrode the nerve root branch was gently lifted using a nerve probe and positioned in the electrode groove. Once the electrode was positioned under the nerve root, the dural sac was released and the miniature electrode securely held the nerve root. The wiring between the mold and amplifier was soft and loose. No micromanipulator was used to hold the electrode. Nerve roots and electrodes were immersed in a pool of mineral oil. All surgical procedures were approved by the institutional animal investigation committee.

Nerve activity recorded from two channels was amplified with an A.C. preamplifier, displayed on an oscilloscope, monitored by an audio recorder, and recorded on an FM tape recorder (MR-30; TEAC, Montebello, CA; Figure 2). The analog tape data were then digitized and analyzed with a computer using the Enhanced Graphics Acquisition and Analysis (EGAA) system (R.C. Electronics Inc., Goleta, CA).

**Determination of Receptors in C5/6 Capsule and Surrounding Muscles.** The C5/6 facet joint gap was determined by palpating the joint space using a blunt glass rod. The joint gap served as a reference marker in identifying and locating various probing locations on the capsule proper. Nine arbitrary probing spots were designated on the dorsal aspect of capsule. They were medial (spot 1), intermediate (spot 2), and lateral (spot 3) spots on the joint gap (Figure 2); medial (spot 4), intermediate (spot 5), and lateral (spot 6) spots on the rostral margin of dorsal capsule; and medial (spot 7), intermediate (spot 8), and lateral (spot 9) spots on the caudal margin of dorsal capsule.

Eight or 15 V electrical stimulus with a pulse duration of 0.1 millisecond at 1 Hz was used to stimulate these spots on the dorsal capsule to provoke the activity of receptors. Classification of evoked units was determined by conduction velocity (CV). CVs were determined by two methods. One method was to measure the distance between stimulated spot and recording site on the cervical dorsal root divided by the unit latency (Figure 3). The unit latency was determined by the duration between electric stimulation artifact and the identified unit. The



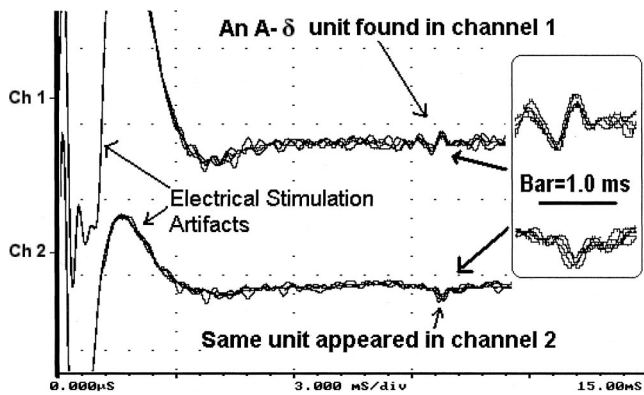


Figure 3. An A- $\delta$  unit evoked by 15 V electrical pulse and recorded in two electrodes. Stimulus artifact and units are seen in both channels (from 0 to 3.6 milliseconds). The unit appeared at 9.4 milliseconds in channel 1 and at 9.7 milliseconds in channel 2. The latency of the unit between the two channels equaled 0.3 milliseconds. The CV of this unit equaled 3.3 m/s (1 mm divided by 0.3 milliseconds). The distance from the stimulated site to the recording site was approximately 32 mm. Using the electrical stimulation method, the CV of this same unit equaled 3.4 m/s (32 mm divided by 9.4 milliseconds).

second method was to divide the distance between the two electrodes by the time difference for the unit to appear in the two recording channels. A- $\beta$  units were usually detected using the second method (Figure 4).

**Responses of Capsule to Mechanical Stimulation.** After electrical stimuli, the dorsal aspect of C5/6 facet joint was probed using a nylon filament attached to a load cell transducer. The signals from the load cell transducer were amplified and sent through an analog-to-digital converter to a computer and recorded together with neural activity. The tip of the nylon filament was broadened into a sphere shape 2 mm in diameter. Force was applied by pressing the nylon filament perpendicularly against the capsule. The probed areas were the same designated 9 spots.

The upper articular process of the left C5/6 facet joint was cut down to pedicle using bone ronguers with the C5/6 facet capsule intact. The capsule was pulled by an actuator (Parker Hannifin Co., Roherk Park, CA) controlled by computer. The

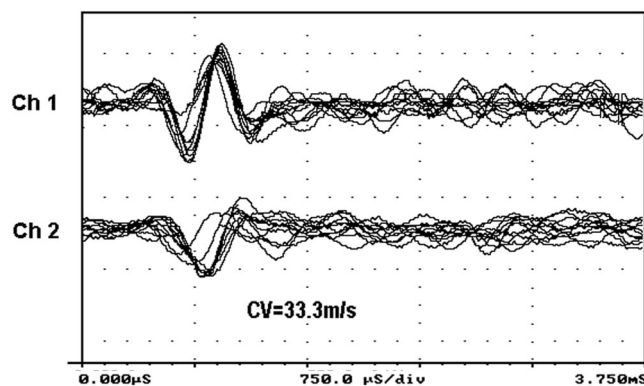


Figure 4. An A- $\beta$  unit recorded using the dual bipolar electrode recording method. The CV equaled 33.3 m/s, based on the 1-mm distance between the two electrodes divided by the latency of this unit between two channels (0.03 milliseconds).

pulling displacement rate was 0.5 mm per second. Loads were recorded from a load cell attached to the actuator. Pulls ranged from 2 to 26 mm in increments of 2 mm until the capsule was torn. Capsular strains were recorded during the load application, and these procedures have been described in detail elsewhere.<sup>15</sup>

**Data Analysis of Neural Activity.** Multiunit discharge analysis: Baseline discharge rate, discharge rate in response to stretch, and discharge rate after the release of stretch (after-discharge) were analyzed using EGAA software. These multiunit discharge rates were determined by setting a voltage threshold above which all nerve discharge spikes were counted.

Single unit discharge analysis: Single unit discharge rate was determined by counting the total number of units whose wave-shape matched the template waveshapes during the period of interest. After a unit was classified based on its CV, the wave-shape of this unit was saved as a template for identifying and counting the same unit in the multiunit discharges. Waveshape recognition software in the EGAA system was used for this procedure. The baseline discharge rate, discharge rate during stretch, and discharge rate after the release of stretch were analyzed for all identified units.

## ■ Results

### *Anatomy of C6 Dorsal Root*

The goat cervical C6 nerve root runs in a fan shape with 5 to 6 branches. The length of individual branches ranged from 2 to 6 mm with a diameter of 0.7 to 0.9 mm within the dural sac. Among all branches of C6 dorsal roots, the third branch of the C6 nerve root was the best for recording facet joint capsule activity. The clearest neural signals could be recorded in this branch.

### *Interaction Between Electrode and Nerve Root*

The miniature electrode assembly securely held the nerve root by being held down between the dorsal root and spinal cord. The wax hook isolated the electrodes from touching soft tissue other than the nerve root. No discernible motion was found between electrodes and the nerve root when the goat neck moved because of pulling on the capsule.

### *Feasibility of Unit Classification*

CV of different units could be obtained by two methods. As shown in Figure 3 in one goat experiment, using the dual-bipolar recording method, the CV equaled 3.33 m/s (1 mm divided by 0.3 milliseconds). Using the electrical-stimulation method, the CV equaled 3.4 m/s (Figure 3).

The bipolar-electrode recording method (Figure 4) was required to find the CV of A- $\beta$  units. A- $\beta$  unit wave-shapes usually appeared within the first 3 milliseconds after electrical stimulation was applied to the capsule.

A- $\delta$  and C-fiber units in these multiunit discharges could be detected using both the electrical-stimulation method and dual-bipolar-electrodes recording method (Figure 4). A- $\beta$  (55 units), A- $\delta$  (58 units), and C-fibers (62 units) have been identified using these methods. Their CVs ranged from 0.5 m/s to 70 m/s.

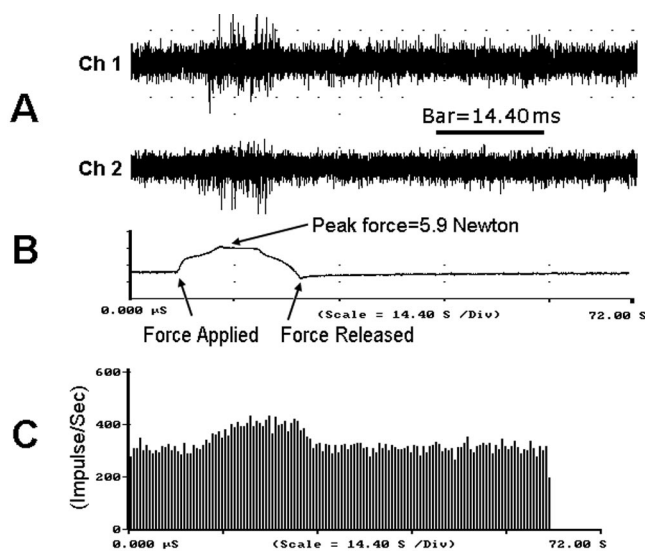


Figure 5. **A**, Neural activity in channel 1 and channel 2. **B**, A force trace recorded simultaneously. **C**, Multiunit discharge rate changed in histogram in response to pulling force on the capsule. Multiunit discharge rate increased as force increased and returned to baseline when force was released.

#### Neural Activity in Response to Compression on the Capsule

Neural activity evoked by compression on the dorsal aspect of capsule was recorded in both channels. Multiunit discharges increased when local compressive force applied with the nylon filament increased and returned to baseline when the force was released. More single units were found in the dorsal lateral capsule region over the joint gap (spot 3) than other spots (*t* test,  $P < 0.05$ ). The dorsal lateral region of the capsule was attached by both the muscles and tendons.

#### Neural Activity in Response to Stretching Capsule

It was found that pulling the capsule via the actuator led to multiunit discharge rate increase (Figure 5). Pulling the capsule  $< 6$  mm, multiunit discharge rate increased during stretch, and returned to the baseline when stretch was released. At higher displacements, more force was required, and increased afterdischarges were produced.

The multiunit discharges included units from all activated receptors including A- $\beta$ , A- $\delta$ , and C-fiber units. Analyzing each single unit individually, it was found that these units responded to stretch of the capsule. Stretch of the capsule led to single unit discharge rate increase with an increase of force (Figure 6). Single unit discharge rate decreased when stretch was released.

Pulling the capsule until a subcatastrophic injury occurred in the capsule led to an increase of afterdischarges. Subcatastrophic injury was defined as a dip in the force trace before reaching maximum force.<sup>7</sup> After pulling the capsule until the capsule was completely broken, both multiunit and single unit afterdischarges decreased. Further stimulation did not evoke any activity in the capsule.

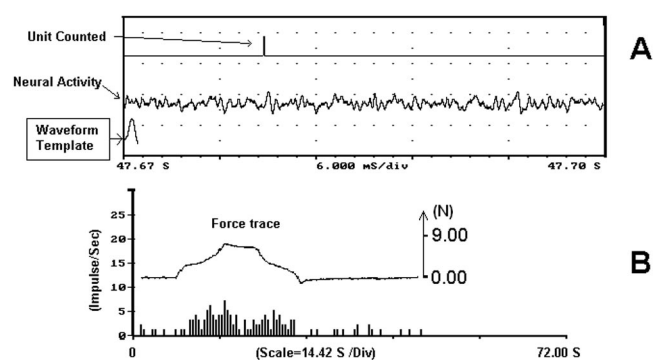


Figure 6. An identified A- $\beta$  unit response to stretch of the capsule using the waveshape recognition method. **A**, When an action potential waveshape matched the template waveshape, this spike was counted as one impulse. **B**, The histogram of the same A- $\beta$  unit over the time when the capsule was pulled for 6 mm. The baseline discharge rate of this A- $\beta$  unit was approximately 1 impulse per second before stretch. When the capsule was pulled by a force of 8.89 N, the discharge rate increased to 5 impulses per second. When force was released, the discharge rate returned to the baseline discharge rate.

#### Discussion

The outcomes of this study demonstrated that there are sensory receptors in the goat cervical facet joint capsule that can respond to mechanical load. This neural activity can be monitored using standard neurophysiology recording equipment, custom-made miniaturized hook electrodes, and waveshape recognition software.

#### Disadvantage of Conventional Bipolar Electrodes

As we found in this study, the goat C6 spinal nerve roots are very short and not conducive to single or dual bipolar electrode recording.<sup>16,17</sup> Conventional electrodes held by a micromanipulator must be fixed on the operating table,<sup>16,17</sup> and the electrodes cannot move along with the goat body. It was found that a force of  $>150$  N was required to tear the capsule. This force would move the goat neck as well as the nerve roots on the electrodes, causing motion artifact in the nerve recordings and damage to the nerve root.

#### Advantage of New Miniature Dual Bipolar Electrodes

With these limitations and the need for small and mobile electrodes, we present here the successful design and development of miniature platinum wire electrodes in a wax platform. The "File-A-Wax" used to fabricate the mold is used worldwide and is ideal for machining and hand curving into any designated shape. In our study, the wax was hard enough to hold two electrodes (4 wires in total) and separated the dorsal root from surrounding soft tissue. The small diameter electrode wires could thus hook and hold the dorsal root securely.

The Teflon coated wires running between the wax bed and the amplifier were tiny and flexible. When the goat body moved, the electrodes moved with no motion between electrode and nerve root. The waveshapes obtained by these miniature electrodes were clear so that conduction velocities of units could be measured. The

response of the facet joint capsule units could be clearly monitored using these miniature electrodes as evidenced by enhanced multiunit and single unit response to increased mechanical stimulation. The recorded neurophysiologic data in this study demonstrated that these specially designed miniature electrodes made it feasible to reliably record neural activity in short cervical dorsal roots.

### **Physiologic Classification and Functions of Units**

Nerve fibers are generally classified into A- $\alpha$ , A- $\beta$ , A- $\delta$ , and C fibers based on their conduction velocity. A- $\beta$  fibers conduct activity from proprioceptors that sense vibration, deep pressure, and touch. A- $\delta$  fibers function to sense deep pressure and pricking pain. C-fibers sense crude touch and pressure and aching pain. CVs of A- $\beta$  fibers range from 30 to 90 m/s, those of A- $\delta$  fibers range from 6 to 30 m/s, and those of C fibers range from 0.5–2 m/s.<sup>18</sup>

### **Comparisons of Two Methods to Obtain CVs**

In this study, the dual-bipolar recording method was used to detect A- $\beta$  units, as the waveshapes of fast-conducting A- $\beta$  units were hidden in the stimulation artifact in the electrical stimulation method. The dual-bipolar recording method as well as electrical stimulation could be used to detect C fiber and A- $\delta$  units (Figure 3), which were not hidden in the stimulation artifact. The CVs obtained using electrical-stimulation equaled CVs obtained using the dual-bipolar method.

### **Sensory Function of Cervical Spine Facet Joint Capsule**

Determination and classification of different units in the capsule are important to understand sensory functions including nociception. After a unit was identified, its waveshape was saved as a template, and its discharge rate was measured in response to stretch of the capsule. This study showed that the waveshape recognition method could search and match the same waveshape in multiunit discharges evoked by mechanical stimulation. Thus, receptor responses to local mechanical compression and capsule stretch could be determined and analyzed.

A- $\beta$ , A- $\delta$ , and C units were found in the dorsal capsule, and they responded to mechanical stimulation. This indicated that the capsule has sensory functions.

When a subcatastrophic injury was produced in capsule, these units appeared to be sensitized because their spontaneous baseline discharge rate increased. When catastrophic injury was produced in the capsule, the discharge rate of these units decreased. This suggests that catastrophic injury can destroy these sensory receptors, leading to lowered sensory input.

### **Significance**

To our knowledge, no previous study on cervical facet joints has been reported that addresses the neurophysiology and biomechanics of cervical facet loading. Knowledge of the magnitude of facet capsule strain that leads to pain is important in understanding the etiology of pain after whiplash.

## **Conclusions**

Custom designed miniature dual bipolar electrodes are suitable for recording neural activity from the cervical spinal dorsal root and for characterizing the receptors in the facet joint capsule. Both electrical-stimulation and dual bipolar electrodes can be used to identify specific units and measure their conduction velocities. Higher conduction velocities can be better detected using the dual bipolar recording electrodes. The responses of receptors in the capsule to mechanical stimulation can be recorded and analyzed. The experimental methodology appears promising for additional studies of the neurophysiology and biomechanics of cervical facet capsules and other soft tissues of the spine. The long-term goal of this work is to provide further insight into whiplash-associated neck pain.

### **Key Point**

- Custom designed miniature dual bipolar electrodes recorded neural activity in goat cervical facet joint. Multiunit and A- $\beta$ , A- $\delta$ , C-fiber units were found to respond to mechanical stimuli.

### **Acknowledgments**

The authors thank Amrita Vempati and Anita Singh for valuable assistance.

## **References**

1. Barnsley L, Lord S, Bogduk N. Whiplash injury. *Pain* 1994;58:283–307.
2. Deng B, Begeman PC, Yang KH, et al. Kinematics of human cadaver cervical spine during low speed rear-end impacts. Proceedings of 44th Stapp Car Crash Conference, 2000 November 6–8; Atlanta, GA. Warrendale, PA: Society of Automotive Engineers, Inc. 2000:171–88.
3. Barnsley L, Lord SM, Wallis BJ, et al. The prevalence of chronic cervical zygapophysial joint pain after whiplash. *Spine* 1995;20:20–5.
4. Ono K, Kaneoka K, Wittek A, et al. Cervical injury mechanism based on the analysis of human cervical vertebral motion and head-necktorso kinematics during low-speed rear impacts. Proceedings of 41st Stapp Car Crash Conference, November 13–14, 1997; Orlando, FL. Warrendale, PA: Society of Automotive Engineers, Inc. 1997:339–56.
5. Siegmund GP, Brault JR, Wheeler JB. The relationship between clinical and kinematic responses from human subject testing in rear-end automobile collisions. *Accid Anal Prev* 2000;32:207–17.
6. Siegmund GP, King DJ, Lawrence JM, et al. Head/neck kinematic response of human subjects in low-speed, rear-end collisions. Proceedings of 41st Stapp Car Crash Conference, November 13–14, 1997; Orlando, FL. Warrendale, PA: Society of Automotive Engineers, Inc. 1997:357–71.
7. Winkelstein BA, Nightingale RW, Richardson WJ, et al. Cervical facet joint mechanics: its application to whiplash injury. Proceedings of 43rd Stapp Car Crash Conference, 1999 October 25–27; San Diego, CA. Society of Automotive Engineers, Inc. 1999:243–52.
8. Yoganandan N, Pintar FA, Cusick JF. Biomechanical analyses of whiplash injuries using an experimental model. *Accid Anal Prev* 2002;34:663–71.
9. Kallakuri S, Singh A, Chen C, et al. Demonstration of substance P, calcitonin gene related peptide and protein gene product 9.5 containing nerve fibers in human cervical facet joint capsules. *Spine* 2004;29:1182–1186.
10. McLain RF. Mechanoreceptor endings in human cervical facet joints. *Spine* 1994;19:495–501.
11. Kandziora F, Pflugmacher R, Scholz M, et al. Comparison between sheep and human cervical spines: an anatomic, radiographic, bone mineral density, and biomechanical study. *Spine* 2001;26:1028–37.
12. Karaićovic EE. Comparison between sheep and human cervical pedicles. *Spine* 2001;26:2640–1.
13. Wilke HJ, Kettler A, Wenger KH, et al. Anatomy of the sheep spine and its comparison to the human spine. *Anat Rec* 1997;247:542–55.



14. Pintar FA, Mayer RG, Yoganandan N, et al. Child neck strength characteristics using an animal model. Proceedings of 44th Stapp Car Crash Conference, 2000 November 6–8; Atlanta, GA. Warrendale, PA: Society of Automotive Engineers, Inc. 2000:77–83.
15. Lu Y, Chen C, Kallakuri S, et al. A method for characterizing the relationship between nerve discharge and mechanical strain in vivo in soft tissue injury. *Er Spine J*. In press. 2005.
16. Chen CY, Cavanaugh JM, Ozaktay AC, et al. The effects of phospholipase A2 on lumbar nerve root structure and function. *Spine* 1997;22:1057–64.
17. Takebayashi T, Cavanaugh JM, Cunejt OA, et al. Effect of nucleus pulposus on the neural activity of dorsal root ganglion. *Spine* 2001;26:940–5.
18. Guyton AC. Sensory receptors and their basic mechanism of action. In: Guyton AC, editor. *Textbook of Medical Physiology*. 6th edition. Philadelphia: WB Saunders Co. 1981:588–96.

# Expressions of Membrane-Type I Matrix Metalloproteinase, Ki-67 Protein, and Type II Collagen by Chondrocytes Migrating from Cartilage Endplate into Nucleus Pulposus in Rat Intervertebral Discs

## A Cartilage Endplate-Fracture Model Using an Intervertebral Disc Organ Culture

Ki-Won Kim, MD,\* Kee-Yong Ha, MD,\* Jong-Beom Park, MD,\* Young-Kyun Woo, MD,\*  
Ha-Na Chung, BS,\* and Howard S. An, MD†

**Study Design.** Immunohistochemistry was performed in organ-cultured intact and cartilage endplate (CE)-fractured rat intervertebral discs (IVDs).

**Objectives.** To demonstrate biologic events associated with migration of chondrocytes from hyaline CE into nucleus pulposus (NP).

**Summary of Background Data.** It was recently revealed that the transition from a notochordal NP to a fibrocartilaginous NP in the rabbit IVD is accomplished exogenously by chondrocytes migrating from CEs into the NP. This observation has not been studied in other animal models, and the biologic events associated with chondrocyte migration have not been elucidated in the literature.

**Methods.** IVDs including cranial and caudal CEs were obtained from 4-week, 6-month, 12-month, and 18-month old Wistar rats. To accelerate chondrocyte migration, CEs of IVDs were fractured and cultured for 48 hours. IVDs without CE-fracture were used as a control for each age group. Expressions of membrane-type I matrix metalloproteinase (MT1-MMP, as a marker for cell migration and extracellular matrix digestion) and Ki-67 protein (as a proliferation marker) and pericellular deposition of type II collagen (as a marker for fibrocartilaginous matrix) by the chondrocytes migrating from CE into NP were examined immunohistochemically.

**Results.** In the control groups, chondrocyte migration limited only along the periphery of the notochordal NP and no chondrocytes were inside the NP proper. However, all the IVDs in the CE-fracture groups showed direct and more extensive migration of chondrocytes from CEs into the NP proper. The migrating chondrocytes in both control and CE-fracture groups expressed MT1-MMP and Ki-67 protein and deposited type II collagen in the NP.

**Conclusions.** This report demonstrates the chondrocyte migration from CE into NP in the organ-cultured rat IVDs. This phenomenon is accelerated in the presence of CE fracture. The chondrocytes migrating from CEs into the NP expressed MT1-MMP and Ki-67 protein and deposited type II collagen. These biologic strategies probably enable chondrocytes of the hyaline CE to migrate into the ectopic NP region, replace notochordal cells, and change the notochordal tissue into fibrocartilage. These results suggest that similar biologic mechanisms may be involved in the natural transition from the notochordal NP to the fibrocartilaginous NP in other animal models, including human.

**Key words:** cartilage endplate, membrane-type 1 matrix metalloproteinase (MT1-MMP), Ki-67 protein, type II collagen, chondrocyte, nucleus pulposus, intervertebral disc, organ culture. **Spine 2005;30:1373–1378**

The nucleus pulposus (NP) of the intervertebral disc (IVD) undergoes a chronological transition from a notochordal NP to a fibrocartilaginous NP.<sup>1–4</sup> This transition accompanies changes in cell types from notochordal cells to chondrocytes and in extracellular matrix (ECM) from highly vacuolated notochordal tissue to fibrocartilage. It was recently demonstrated that these transitional changes in the intact rabbit IVD are accomplished exogenously by chondrocytes migrating from hyaline cartilage endplates (CEs) into the notochordal NP.<sup>5</sup> These chondrocytes deposit fibrocartilaginous lamellas and fibers in a centripetal manner.<sup>5</sup> During the transition, notochordal cells gradually disappear and chondrocytes take the place of notochordal cells.<sup>1</sup>

Matrix metalloproteinases (MMPs), also known as matrixin, are a subfamily of the metzincin superfamily.<sup>6,7</sup> To date, 24 different vertebrate MMPs have been identified.<sup>8</sup> Among them, 6 MMPs have membrane-anchor signals and are referred to as membrane-type MMPs (MT-MMPs).<sup>8,9</sup> MT1-MMP was the first to be identified,<sup>9,10</sup> and its substrates include proteoglycans, collagens type I, II, and III, CD44, gelatin, fibronectin, and vitronectin.<sup>11</sup> The enzyme also activates other proMMPs, such as proMMP-2 (gelatinase A) and proMMP13 (collagenase 3).<sup>11</sup> MT1-MMP localizes at the front of migrating cells and promotes cell migration by digesting pericellular ECM barriers, triggering MMP

From the Department of \*Orthopedic Surgery St. Mary's Hospital, The Catholic University of Korea Seoul, Korea; and the †Department of Orthopedic Surgery Rush-Presbyterian-St. Luke Medical Center Chicago, Illinois.

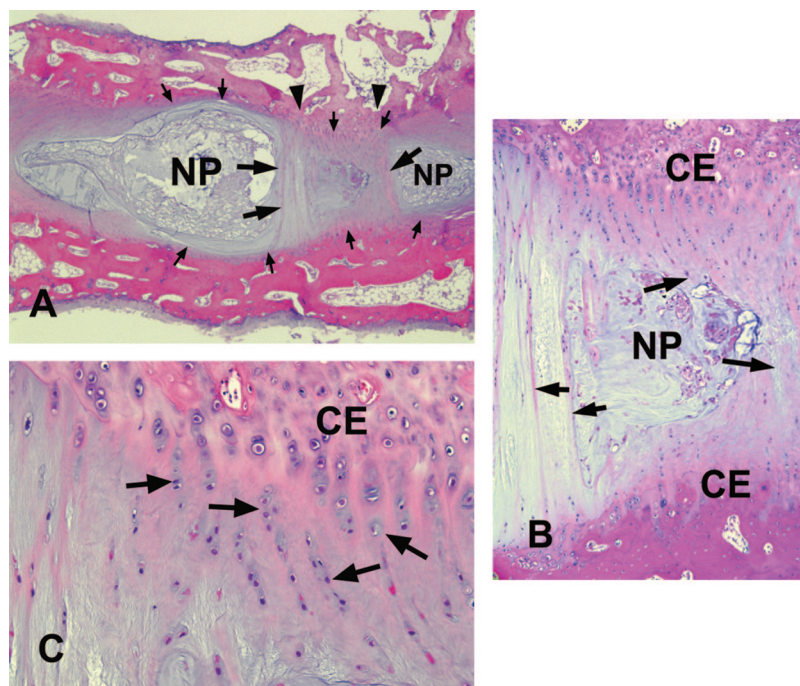
Acknowledgment date: March 29, 2004. First revision date: May 6, 2004. Acceptance date: July 20, 2004.

The manuscript submitted does not contain information about medical device(s)/drug(s).

No funds were received in support of this work. No benefits in any form have been or will be received from a commercial party related directly or indirectly to the subject of this manuscript.

Address correspondence to Ki-Won Kim, MD, Department of Orthopedic Surgery, St. Mary's Hospital, The Catholic University of Korea, 62 Yoido-dong, Youngdeungpo-ku, Seoul, 150-713, Korea; E-mail: kiwonk@lycos.co.kr

Figure 1. H&E stain of organ cultured 6-month-old rabbit intervertebral disc with CE fracture. **A**, Invasions of fibrocartilaginous fibers (large arrows) from both cranial and caudal CEs into the notochordal NP ( $\times 40$ ). Small arrows and arrowheads indicate CEs and fracture sites, respectively. **B**, The invading fibrocartilaginous fibers (arrows) divide the notochordal NP, entrapping notochordal tissue between them ( $\times 100$ ). **C**, The fibrocartilaginous fibers are formed by chondrocytes (arrows) migrating from CEs ( $\times 400$ ). CE, cartilage endplate; NP, nucleus pulposus.



activation cascades and shedding or modulating cell adhesion molecules.<sup>9,11-13</sup>

It is unknown how the chondrocytes in the hyaline CE are able to migrate into the ectopic NP region, replace notochordal cells, and change the highly vacuolated notochordal tissues into fibrocartilaginous matrix. These chondrocytes must remove the ECM barriers and gain motility to reach the ectopic NP region. Also, cell proliferation must take place in order to replace the notochordal cells, and deposition of type II collagen must follow to form the fibrocartilaginous matrix.

In this study, using an organ culture technique and a CE-fracture model,<sup>14-16</sup> expressions of MT1-MMP (a marker for cell migration and ECM digestion) and Ki-67 protein (a proliferation marker) and pericellular deposition of type II collagen (a marker for fibrocartilaginous matrix) by the chondrocytes migrating from hyaline CEs into the notochordal NP were observed immunohistochemically. The authors used the CE-fracture model based on our preliminary experiment showing acceleration of the chondrocyte migration from the CE into the notochordal NP by CE fractures.

## Materials and Methods

**Tissue Processing and Organ Culture.** Lumbar spinal column (L1-L6) was removed under sterile conditions from 4-week, 6-month, 12-month, and 18-month-old male Wistar rats with two rats for each age group. Five lumbar IVDs (from L1-L2-L5-L6), including both cranial and caudal CEs, were harvested from the spinal column. Our preliminary results from the organ culture of rabbit IVDs showed that CE fracture accelerated chondrocyte migration from CEs into NP (Figure 1). Based on these results, CEs of six IVDs for each age group were fractured with surgical hemostat, which resulted in herniation of some NP materials at the fracture sites. Four IVDs without CE-fracture were used as a control for each age group.

IVDs were placed into multiwell tissue culture plates and cultured in Dulbecco's modified Eagle's medium/high glucose (Invitrogen, Carlsbad, CA), supplemented with 30% fetal bovine serum (Invitrogen), vitamin K ( $10^{-4}$  mol/L, Sigma, St. Louis, MO), ascorbic acid (50  $\mu$ g/mL; Sigma), 1% penicillin-streptomycin (Invitrogen). IVDs were maintained for 48 hours at 37°C in a humidified atmosphere of 5% CO<sub>2</sub>. IVDs were then fixed in 4% para-formaldehyde at 4°C for 24 hours and decalcified in 20% EDTA (pH 7.4) for 2 to 6 weeks. The IVD samples then were dehydrated and embedded in paraffin. From the resulting blocks, 3  $\mu$ m-thick midsagittal sections of the IVDs were cut and mounted on poly-L-lysine-coated slides for immunohistochemistry.

**Immunohistochemistry.** For immunohistochemical staining, monoclonal antibodies were purchased (MT1-MMP, Chemicon, Temecula, CA; Ki-67 protein and type II collagen, Labvision, Fremont, CA). The antibody against type II collagen specifically recognizes the triple helical domain. The paraffin sections were deparaffinized in xylene and rehydrated in graded alcohol (100%, 90%, 80%, and 70%). The endogenous peroxidase was subsequently blocked by 0.3% H<sub>2</sub>O<sub>2</sub> for 30 minutes. The sections were boiled in 10% citrate buffer (pH 6.0) for 15 minutes for Ki-67 and MT1-MMP or treated with pepsin and incubated at 37°C for 2 hours for type II collagen. The sections were then incubated with relevant primary antibodies at 4°C for 16 hours. The sections were then exposed to a streptavidin-biotin-peroxidase complex, and color was developed with 3,3'-diaminobenzidine hydrochloride. Mayer's hematoxyline was used for counterstaining.

## Results

In the control groups, there were no chondrocytes in the NP proper. Chondrocytes that originated from CEs encircled the periphery of the notochordal NP (or the inner border of the inner annulus fibrosus), which is somewhat similar to the findings observed in the rabbit IVD (Figure



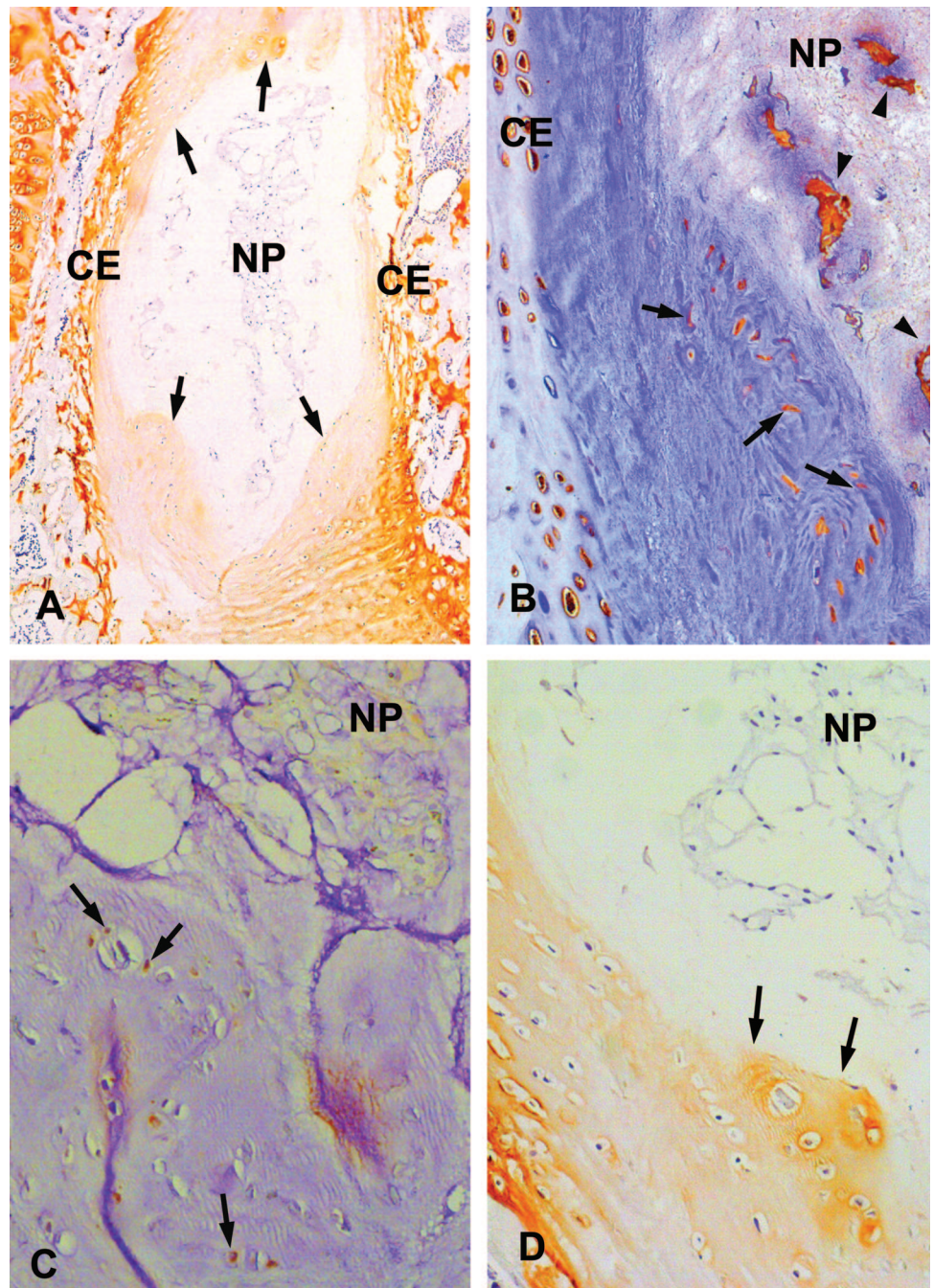


Figure 2. **A**, Immunohistochemical staining of type II collagen in 12-month-old rat intervertebral disc without CE fracture shows limited chondrocyte migration and fibrocartilage formation along the periphery of the notochordal NP (arrows,  $\times 40$ ). **B to D**, A higher magnification of the left bottom arrow region in Figure 2A. Some of these chondrocytes express MT1-MMP (**B**, arrows,  $\times 200$ ) and Ki-67 protein (**C**, arrows,  $\times 400$ ) and all of them deposited type II collagen along the periphery of the notochordal NP (**D**,  $\times 400$ ). Notochordal cells also express MT1-MMP (**B**, arrow heads). There is no pericellular deposition of type II collagen around notochordal cells (**D**). CE, cartilage endplate; NP, nucleus pulposus.

2A).<sup>5</sup> Some of these chondrocytes expressed MT1-MMP (Figure 2B) and Ki-67 protein (Figure 2C) and all of them deposited type II collagen (Figure 2D). Notochordal cells extensively expressed MT1-MMP (Figure 2B). At a microscopic level, there was no visible pericellular deposition of type II collagen around notochordal cells (Figure 2D).

In the CE-fractured groups, direct migration of chondrocytes from CEs into the notochordal NP proper was observed in all CE-fractured IVDs (Figure 3A). There were no appreciable differences among different age groups. In contrast to intact IVDs where chondrocyte migration was limited only along the periphery of the notochordal NP (or the inner border of the inner annulus

fibrosus), chondrocytes in the CE-fractured groups migrated more extensively into the NP proper. Many of chondrocytes that migrated from CEs into the NP expressed MT1-MMP (Figure 3B) and Ki-67 (Figure 3C), and all of them deposited type II collagen. (Figures three-dimensional)

#### ■ Discussion

It has long been a matter of mystery how the gel-like, highly vacuolated notochordal NP undergoes a chronological transition to the fibrocartilaginous NP.<sup>1,3</sup> This transition was assumed to be an endogenous development by cells that reside in the NP<sup>17,18</sup> or an exogenous replacement by cells that exist outside the NP.<sup>3,4,19</sup> It



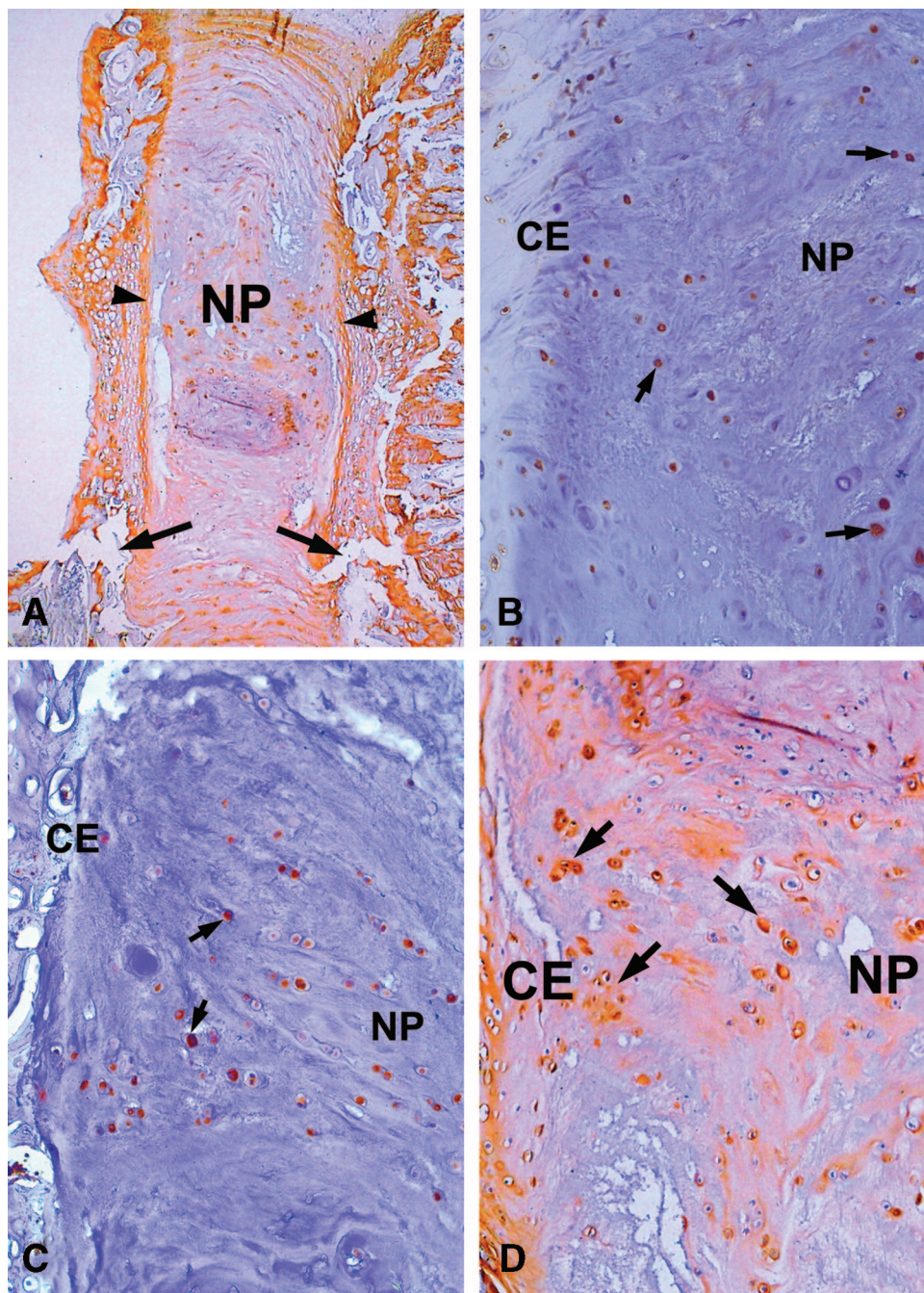


Figure 3. CE-fractured 18-month-old rat intervertebral disc. **A**, Immunohistochemical staining of type II collagen ( $\times 40$ ) shows direct and more extensive chondrocyte migration from CEs (arrowheads) into the NP proper compared with control specimen (Figure 2A). Arrows indicate fracture sites. **B to D**, In this specimen, many of these chondrocytes expressed MT1-MMP (**B**, arrows,  $\times 200$ ) and Ki-67 protein (**C**, arrows,  $\times 400$ ), and all of them deposited type II collagen (**D**, arrows,  $\times 200$ ) in the NP proper. CE, cartilage endplate; NP nucleus pulposus.

was revealed recently that in the rabbit IVD, the transition is accomplished exogenously by chondrocytes migrating from hyaline CEs into the notochordal NP.<sup>5</sup> These chondrocytes deposit fibrocartilaginous lamellas and fibers in a centripetal manner, which decreases the size of the notochordal NP while increasing the lamellar structure of the IVD.<sup>5</sup> However, it is unknown how the chondrocytes of the hyaline CE are able to migrate into the ectopic NP region, change the highly vacuolated notochordal tissues into fibrocartilaginous matrix, and take the place of notochordal cells.

During cell migration, CD44 localizes at the migration front, and this localization is mediated by the cytoplasmic domain that associates with actin cytoskele-

ton.<sup>11,13</sup> By forming a complex with CD44, MT1-MMP also localizes at the migration front.<sup>13</sup> The hemopexin-like domain of MT1-MMP is the site for the complex formation with CD44 and determines its localization on the cell surface.<sup>13</sup> MT1-MMP promotes cell migration by digesting pericellular ECM barriers, triggering MMP activation cascades and shedding or modulating cell adhesion molecules.<sup>9,11-13</sup> In the present study, chondrocytes migrating into the notochordal NP expressed MT1-MMP. MT1-MMP and other MMPs activated by this enzyme probably enable chondrocytes of the hyaline CE to digest both the surrounding hyaline cartilage and the vacuole-rich notochordal tissue, thereby opening a migration pathway from CEs to the notochordal NP. Be-

sides this enzymatic digestion of the ECM barriers, MT1-MMP, as observed in other cells, may also provide motility to the chondrocytes by forming a complex with CD44 and detaching the chondrocytes from the ECM.<sup>9,12,13</sup> This result implies that MT1-MMP expressed by chondrocytes plays an important role in the chondrocyte migration from hyaline CEs into the ectopic NP region, probably by digesting ECM barriers and providing motility.

This study using the antibody that specifically recognizes the triple-helical domain of type II collagen showed that pericellular deposition of type II collagen was visible in the NP only around migrating chondrocytes but not around notochordal cells. Because the triple helical domain is contained in both type IIA and IIB collagens synthesized by embryonic notochord cells and chondrocytes, respectively, some possibilities can be speculated for the negative staining of the triple-helical domain of type II collagen around notochordal cells. As in the embryonic development of human IVDs, postnatal notochordal cells may synthesize type IIA procollagen, and only the NH<sub>2</sub>-propeptide devoid of the triple helical fibrillar domain may be deposited in the notochordal NP.<sup>20</sup> Alternatively, extracellular proteolysis and/or subsequent self-assembly of notochordal type IIA collagen may differ from those of chondrocytic type IIB collagen, resulting in a type IIA collagen fibril thickness not detectable by microscopy.<sup>21,22</sup> Additional research is needed to elucidate the precise form of type II collagen synthesized by notochordal cells.

During and after migration into the ectopic NP region, the chondrocytes, in terms of deposition of type II collagen, seem to maintain their phenotype. However, the chondrocytes did not form hyaline cartilage but new fibrocartilage in the NP. The formation of the fibrocartilaginous matrix in the NP is thus considered to be processes influenced by numerous factors besides the chondrocytic phenotype.

Ki-67 protein is present in the nuclei of cells in the G<sub>1</sub>, S, and G<sub>2</sub> phases of the cell division cycles as well as in mitosis.<sup>23</sup> Quiescent or resting cells in the G<sub>0</sub> phase do not express the Ki-67 protein.<sup>23</sup> These facts make the Ki-67 protein an excellent marker to determine the growth fraction of a given cell population.<sup>23–25</sup> In intact rabbit IVDs, migration of chondrocytes led to a depletion of chondrocytes in the CE, a disappearance of zonal arrangement of the CE, and a subsequent narrowing of the CE.<sup>5</sup> In the present study, some of chondrocytes in the control groups that migrated along the periphery of the notochordal NP and many of the chondrocytes in the CE-fractured groups that directly migrated into the NP proper were intranuclear Ki-67-positive. This result indicates that chondrocytes not only migrate into the notochordal NP but also proliferate, thereby occupying the NP region. This result also supports our previous idea that chondrocytes of the hyaline CE are the source of chondrocytes of the NP.<sup>5</sup>

This study demonstrated that CE-fracture accelerates chondrocyte migration from CEs into the NP proper, and this phenomenon is more disorganized as compared with the centripetal manner in the intact rat IVDs without CE-fracture. This observation is consistent with our preliminary data in the rabbit model in that more chondrocytes were observed in the NP earlier in the culture period in the presence of CE-fracture (Figure 1, unpublished data). Such acceleration and disorganization of the chondrocyte migration caused by CE-fracture is non-physiologic, leading to an early formation of the disorganized fibrocartilaginous NP. This, in turn, may eventually result in an early pathologic degeneration of the NP. The main purpose of this study, however, was to observe biologic events associated with chondrocyte migration, and therefore, CE fractures were created to accelerate this process and to shorten the culture period to 48 hours.

This study shows that the transition from notochordal NP to fibrocartilaginous NP in the rat NP is the result of biologic events associated with chondrocyte migration from hyaline CEs into the notochordal NP. As in rabbit,<sup>5</sup> this exogenous replacement of notochordal cells and tissues with chondrocytes and fibrocartilage is an unavoidable natural course of the rat NP. In terms of exogenous replacement, these transitional changes are not degenerative changes of the NP, although they can be the starting point of the degeneration. It is our view that degeneration probably develops later by numerous factors affecting functions or apoptosis of the NP chondrocytes following the completion of the transition (i.e., the formation of the fibrocartilaginous NP).

There are limitations to this study. The organ culture technique used in this study has been developed recently,<sup>14,15</sup> and more basic studies are still required to understand the cellular and biochemical processes in this model.<sup>16</sup> Previous studies showed that this method can maintain cell functions by 1 week in the rat IVD<sup>15</sup> and over 2 weeks in the rabbit IVD.<sup>16</sup> In the present study, all immunohistochemical analyses were performed at 48 hours, which appears safe to draw our conclusions. In addition, our CE-fracture model does not represent the natural transition occurring slowly with age. However, our control specimens showed similar expression patterns of MT1-MMP, Ki-67, and type II collagen by the chondrocytes migrating along the periphery of the notochordal NP, although their expressions were less extensive than in the CE-fractured specimens.

This study demonstrates that chondrocytes migrating from CE into NP expressed MT1-MMP and Ki-67 protein and deposited type II collagen. These biologic strategies probably enable chondrocytes of the hyaline CE to migrate into the ectopic NP region, replace notochordal cells, and change notochordal tissue into fibrocartilage. The results suggest that similar biologic mechanisms may be involved in the natural transition from the noto-



chordal NP to the fibrocartilaginous NP in other species, including those transitions in human subjects.

### ■ Key Points

- Expressions of MT1-MMP, Ki-67 protein, and type II collagen by the chondrocytes migrating from cartilage endplate (CE) into nucleus pulposus (NP) were examined immunohistochemically in organ-cultured intact and CE-fractured rat intervertebral discs.
- The chondrocytes migrating from CEs into the NP expressed MT1-MMP and Ki-67 protein and deposited type II collagen.
- These biologic strategies probably enable chondrocytes of the hyaline CE to migrate into the ectopic NP region, replace notochordal cells, and change the notochordal tissue into fibrocartilage.

### Acknowledgments

As described by Hunter CJ, Matyas JR, Duncan NA (*Tissue Eng* 9:667–677, 2003), the authors use the term “notochordal cell” to mean cells derived from the notochord, as opposed to notochord cell, which would imply that the cells were identical to those found in the embryonic notochord.

### References

1. Buckwalter JA. Aging and degeneration of the human intervertebral disc. *Spine* 1995;20:1307–14.
2. Humzah MD, Soames RW. Human intervertebral disc: structure and function. *Anat Rec* 1988;220:337–56.
3. Peacock A. Observations on the postnatal structure of the intervertebral disc in man. *J Anat* 1952;86:162–79.
4. Rufai A, Benjamin M, Ralphs JR. The development of fibrocartilage in the rat intervertebral disc. *Anat Embryol (Berl)* 1995;192:53–62.
5. Kim KW, Lim TH, Kim JG, et al. The origin of chondrocytes in the nucleus pulposus and histologic findings associated with the transition of a notochordal nucleus pulposus to a fibrocartilaginous nucleus pulposus in intact rabbit intervertebral discs. *Spine* 2003;28:982–990.
6. Egeblad M, Werb Z. New functions for the matrix metalloproteinases in cancer progression. *Nat Rev Cancer* 2002;2:161–74.
7. Nagase H, Woessner JF Jr. Matrix metalloproteinases. *J Biol Chem* 1999;274:21491–4.
8. Visse R, Nagase H. Matrix metalloproteinases and tissue inhibitors of metalloproteinases: structure, function, and biochemistry. *Circ Res* 2003;92:827–39.
9. Seiki M. The cell surface: the stage for matrix metalloproteinase regulation of migration. *Curr Opin Cell Biol* 2002;14:624–32.
10. Sato H, Takino T, Okada Y, et al. A matrix metalloproteinase expressed on the surface of invasive tumour cells. *Nature* 1994;370:61–5.
11. Seiki M. Membrane-type 1 matrix metalloproteinase: a key enzyme for tumor invasion. *Cancer Lett* 2003;194:1–11.
12. Kajita M, Itoh Y, Chiba T, et al. Membrane-type 1 matrix metalloproteinase cleaves CD44 and promotes cell migration. *J Cell Biol* 2001;153:893–904.
13. Mori H, Tomari T, Koshikawa N, et al. CD44 directs membrane-type 1 matrix metalloproteinase to lamellipodia by associating with its hemopexin-like domain. *Embo J* 2002;21:3949–59.
14. Ariga K, Yonenobu K, Nakase T, et al. Mechanical stress-induced apoptosis of endplate chondrocytes in organ-cultured mouse intervertebral discs: an ex vivo study. *Spine* 2003;28:1528–33.
15. Risbud MV, Izzo MW, Adams CS, et al. An organ culture system for the study of the nucleus pulposus: description of the system and evaluation of the cells. *Spine* 2003;28:2652–8.
16. An HS. Point of view. *Spine* 2003;28:2658–9.
17. Trout JJ, Buckwalter JA, Moore KC. Ultrastructure of the human intervertebral disc: II. Cells of the nucleus pulposus. *Anat Rec* 1982;204:307–14.
18. Trout JJ, Buckwalter JA, Moore KC, et al. Ultrastructure of the human intervertebral disc. I. Changes in notochordal cells with age. *Tissue Cell* 1982;14:359–69.
19. Higuchi M, Abe K, Kaneda K. Changes in the nucleus pulposus of the intervertebral disc in bipedal mice. A light and electron microscopic study. *Clin Orthop* 1983;251–7.
20. Zhu Y, McAlinden A, Sandell LJ. Type IIA procollagen in development of the human intervertebral disc: regulated expression of the NH(2)-propeptide by enzymic processing reveals a unique developmental pathway. *Dev Dyn* 2001;220:350–62.
21. Fukui N, McAlinden A, Zhu Y, et al. Processing of type II procollagen amino propeptide by matrix metalloproteinases. *J Biol Chem* 2002;277:2193–201.
22. Gelse K, Poschl E, Aigner T. Collagens—structure, function, and biosynthesis. *Adv Drug Deliv Rev* 2003;55:1531–46.
23. Scholzen T, Gerdes J. The Ki-67 protein: from the known and the unknown. *J Cell Physiol* 2000;182:311–22.
24. Bonfil RD, Gonzalez AD, Siguelboim D, et al. Immunohistochemical analysis of Ki-67, p21waf1/cip1 and apoptosis in marker lesions from patients with superficial bladder tumours treated with vinorelbine intravesical therapy in a preliminary phase I trial. *BJU Int* 2001;88:425–31.
25. Heenen M, Thiriar S, Noel JC, et al. Ki-67 immunostaining of normal human epidermis: comparison with 3H-thymidine labelling and PCNA immunostaining. *Dermatology* 1998;197:123–6.

# Developmental Anomalies of the Cervical Spine in Patients With Fibrodysplasia Ossificans Progressiva Are Distinctly Different From Those in Patients With Klippel-Feil Syndrome

## Clues From the BMP Signaling Pathway

Alyssa A. Schaffer, MD,\*† Frederick S. Kaplan, MD,‡ Michael R. Tracy, MS, MD,\*†  
Megan L. O'Brien, BA,\*† John P. Dormans, MD,\*‡ Eileen M. Shore, PhD,‡  
Richard M. Harland, PhD,¶ and Kenro Kusumi, PhD\*†§

**Study Design.** A radiographic analysis of the cervical spine of 70 patients diagnosed with fibrodysplasia ossificans progressiva (FOP) and 33 diagnosed with Klippel-Feil (KF) syndrome was conducted.

**Objectives.** The objectives of this study were to describe cervical spine abnormalities in patients with FOP, to compare and contrast those findings with the malformations in patients with KF syndrome, and to examine the possible etiology of these abnormalities.

**Summary of Background Data.** Congenital features of diseases often provide seminal clues to underlying etiology and developmental pathways. While progressive metamorphosis of connective tissue to heterotopic bone is the most dramatic and disabling feature of FOP, less severe congenital anomalies of the skeleton are also present. Vertebral fusions observed in KF are consistent with defects in embryonic segmentation.

**Methods.** The cervical spine plain films of 70 FOP patients and 33 KF patients with documented congenital abnormalities were reviewed.

**Results.** Generalized neck stiffness and decreased range of motion were noted in most children with FOP. In the FOP patient group, characteristic anomalies, including large posterior elements, tall narrow vertebral bodies,

and fusion of the facet joints between C2 and C7, were observed. Most notably, these characteristic anomalies of the cervical spine in patients with FOP were distinctly different from those of 33 patients with KF that were examined but were strikingly similar to those seen in mice with homozygous deletions of the gene-encoding noggin, a bone morphogenetic protein (BMP) antagonist.

**Conclusions.** FOP patients exhibit a characteristic set of congenital spine malformations. While the noggin gene (*NOG*) is not mutated in patients who have FOP, these findings extend a growing body of evidence implicating overactivity of the BMP signaling pathway in the molecular pathogenesis of FOP.

**Key words:** fibrodysplasia ossificans progressiva, Klippel-Feil, cervical spine, heterotopic ossification, bone morphogenetic protein pathway, noggin. **Spine 2005;30:1379–1385**

From the Divisions of \*Orthopaedic Surgery and †Human Genetics, Children's Hospital of Philadelphia (Institution B), Philadelphia, PA; Departments of ‡Orthopaedic Surgery and §Pediatrics, University of Pennsylvania School of Medicine (Institution A), Philadelphia, PA; and ¶Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, CA.

Acknowledgment date: April 22, 2004. First revision date: July 28, 2004. Acceptance date: July 29, 2004.

Supported in part by the International FOP Association, the Cali FOP Endowment, the Weldon FOP Endowment, the Betty Laue Resource Center, the Isaac and Rose Nassau Professorship of Orthopaedic Molecular Medicine, and the National Institutes of Health (NIH: 2R01-AR041916) to F.S.K. and E.M.S. This work was also supported by grants to K.K. and J.P.D. from the Cervical Spine Research Society and the Ethel Brown Forrester Fund for Excellence. K.K. is a Hitchings-Elion Fellow of the Burroughs Wellcome Fund, and a Florence R.C. Murray Fellow.

The manuscript submitted does not contain information about medical device(s)/drug(s).

Federal, Foundation, and Professional Organizational funds were received in support of this work. No benefits in any form have been or will be received from a commercial party related directly or indirectly to the subject of this manuscript.

Address correspondence and reprint requests to Frederick S. Kaplan, MD, Isaac and Rose Nassau Professor of Orthopaedic Molecular Medicine, Hospital of the University of Pennsylvania, Department of Orthopaedic Surgery, Silverstein Two–3400 Spruce Street, Philadelphia, PA 19104; E-mail: Frederick.Kaplan@uphs.upenn.edu

Fibrodysplasia ossificans progressiva (FOP) is a rare, autosomal dominant disorder of connective tissue characterized by congenital malformation of the great toes<sup>1–11</sup> and by progressive, disabling heterotopic ossification of soft connective tissue in characteristic anatomic patterns.<sup>1–11</sup> Reproductive fitness is low in patients with FOP due to severe progressive disability. There are few known familial occurrences with most cases occurring sporadically.<sup>9–11</sup>

In addition to great toe malformations, skeletal (or-thotopic) abnormalities in FOP patients also occur in the cervical spine, one of the first areas involved in heterotopic ossification (Figure 1).<sup>1,2,4,6–11</sup> Kaplan *et al* suggested that a diagnosis of FOP can be excluded as a cause of heterotopic ossification in a patient by the absence of malformation of the great toes and by the presence of a normal cervical spine as seen by radiograph, highlighting the significant association between cervical spine abnormalities and FOP.<sup>5</sup> In his review of 28 subjects, Smith observed small vertebral bodies and frequent fusion of large lateral masses, which were unrelated to ossification in the overlying muscles.<sup>9</sup> He proposed that, while fusion of the lateral masses could be present at birth, it was clear that fusion continued in a progressive manner after birth.<sup>9</sup>

We were stimulated to perform this study since one of us (F.S.K.) observed neck stiffness and decreased range of

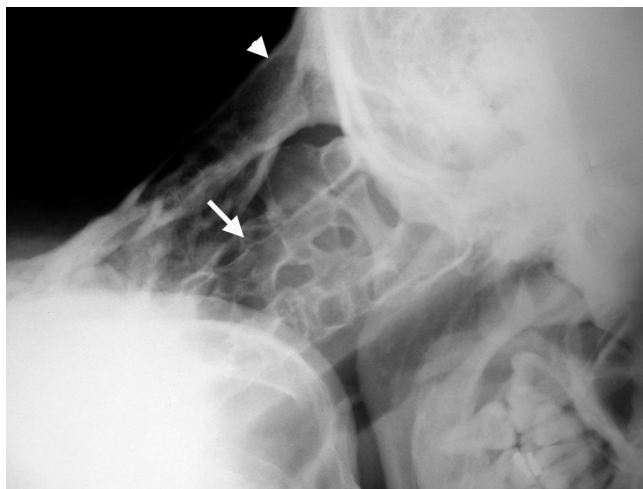


Figure 1. A typical patient with fibrodysplasia ossificans progressiva. Note that the patient has both orthotopic (white arrow) and heterotopic (white arrowhead) bone formation.

motion of the cervical spine in most children with FOP before the formation of heterotopic bone at that anatomic site. Because of their cervical spine abnormalities, FOP patients may incorrectly be diagnosed with Klippel-Feil syndrome (KF), a syndrome characterized by cervical vertebral fusion with short neck, low posterior hairline, and decreased range of motion in the cervical spine.<sup>12–15</sup>

While abnormalities of the cervical spine frequently have been noted as a feature of FOP,<sup>2,4,6,11</sup> these reports focused on a descriptive analysis of the cervical spine abnormalities, and to our knowledge no studies have examined the possible etiology of these malformations. The goal of this study was not only to describe the cervical spine abnormalities of FOP patients but also to propose possible developmental pathways involved in the etiology of the observed abnormalities. Furthermore, we will document how the cervical spine of FOP patients compares and contrasts with that of patients who have KF.

## ■ Materials and Methods

In accordance with an Institutional Review Board approved protocol, available plain films of the cervical spine of 70 patients who had been diagnosed with FOP and referred to Institution A were reviewed. Fifty-three patients had only a single lateral cervical spine film available for review, while 17 patients had a series (two or more films greater than 1 year apart) of lateral cervical spine films available for review. Thirty-seven of the patients were female and 33 were male. Patient age at the time of the first or only radiograph ranged from 7 months to 58 years.

We also reviewed posteroanterior and lateral cervical spine radiographs as well as clinical records of 33 patients diagnosed with KF who were referred to Institution B for evaluation. Nineteen of these patients were female and 14 were male. The KF patients ranged in age from 3 to 22 years.

Individual radiographs and series radiographs for KF and FOP patients were reviewed, noting the age of the patient at the time the radiograph was taken, the shape and size of the verte-

bral bodies, pedicles, and spinous processes, and the presence or absence of congenital anomalies including fusion of the occiput to the first cervical vertebra (occipitalization). The presence, degree, and location of any fusions were noted. In those patients with serial radiographs, fusion progression, involvement of the vertebral bodies in the fusion, and change of disc space height over time were noted.

*Dll3<sup>pm</sup>* and *Nog* mouse mutant lines were maintained in accordance with protocols presented to appropriate Institutional Animal Care and Use Committees. Generation and skeletal characterization of these mouse mutants have been described previously.<sup>16–20</sup> In brief, skeletal preparations of neonates were carried out using Alizarin Red for bone and Alcian Blue for cartilage (*Dll3<sup>pm</sup>*) or Alcian Blue stain alone (*Nog*).

## ■ Results

Of the 70 FOP patients reviewed, all (100%) had orthotopic abnormalities of the cervical spine in addition to heterotopic ossification of the posterior soft tissues. The cervical spine in all (100%) of the FOP patients consisted of narrow vertebral bodies, enlarged pedicles, and large spinous processes. Sixty-three of 70 patients (90%) showed some degree of fusion of the posterior elements. The 7 patients who did not show fusion were less than 1 year of age. None (0%) of the patients had fusion of the occiput to C2 (occipitalization), despite the fact that most had extensive fusion of the cervical spine.

Of the 17 FOP patients with series radiographs, 15 (88%) showed progressive fusion. The narrow vertebral bodies seen in these patients remained separate even though extensive fusion existed in the posterior elements. Only 1 of the 17 patients (6%) with series films showed fusion of the vertebral bodies. Although the narrow vertebral bodies generally remained separate, a progressive narrowing of disc spaces was noted in 11 of the 17 (65%) patients. Progressive fusion of the enlarged lateral masses and spinous processes not associated with heterotopic ossification was common among FOP patients. Fusion between the lateral masses and spinous processes preceded heterotopic ossification (Figure 2). A plain radiograph of the cervical spine in a 2-year-old child with FOP (Figure 2A) showed the presence of small (narrow) vertebral bodies and enlarged lateral masses and spinous processes. There was no obvious fusion between the lateral masses or spinous processes. As the patient increased in age, fusion of the lateral masses and the spinous processes became more prominent on plain radiographs (Figure 2B–D), although no fusion was noted between the occiput and C2. By the time the patient was 8 years old (Figure 2C), a very consistent pattern of fusion was seen at each level of the cervical spine; fusion was present between the lateral masses and the spinous processes but absent between the vertebral bodies. This very regular pattern of fusion was seen in the majority (72%) of the FOP patients examined.

The KF patients did not show the progressive fusion that was seen in the FOP patients. Figure 2 clearly demonstrates the progressive nature of the fusion seen in an



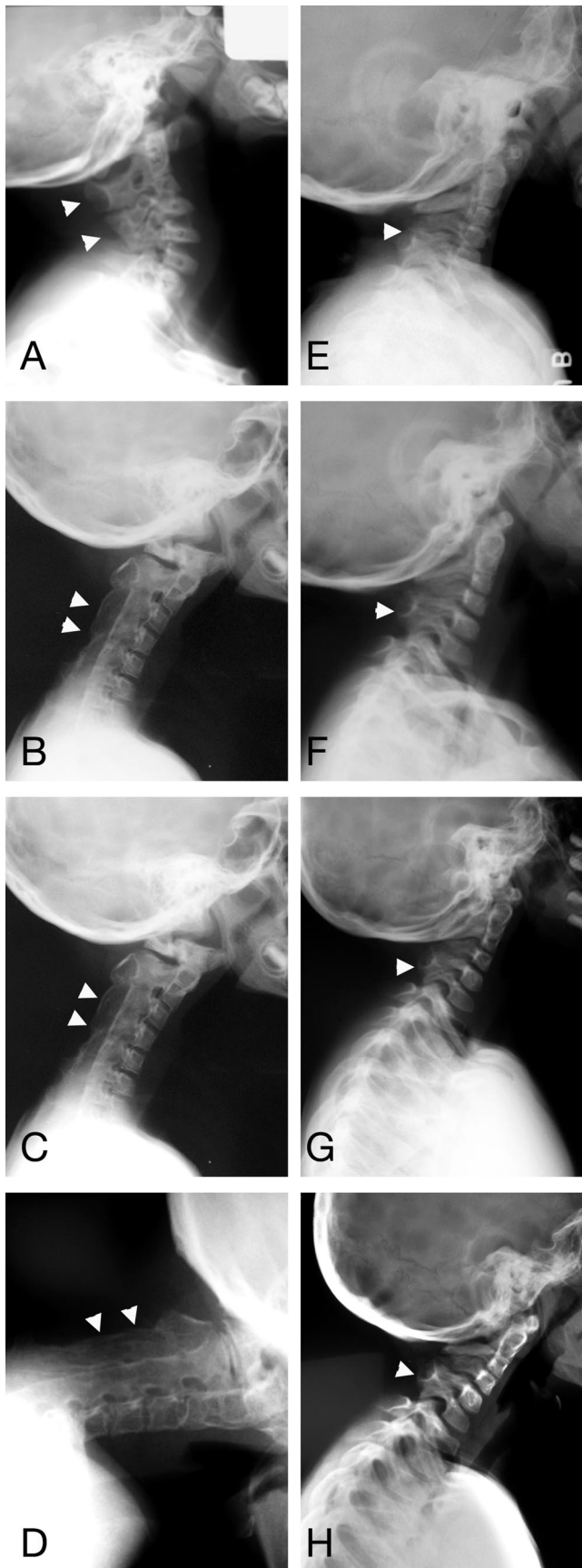


Figure 2. Progressive fusion of the lateral masses and spinous processes in a typical patient with FOP compared to the nonprogressive fusion of a typical patient with KF syndrome over a similar time span. **A–D**, Progressive fusion of the lateral masses and

FOP patient (**A–D**) as compared with a KF patient (**E–H**) who shows no progression of cervical spine ankylosis (fusion). Eight of the 33 KF patients (24%) had single defects (fusion of two vertebrae) of the cervical spine, and 25 (76%) had multiple defects (fusion of more than two contiguous vertebrae). Of the 25 KF patients who had multiple defects, 8 (32%) had major defects involving mainly the posterior elements, 9 (36%) had major defects involving mainly the vertebral bodies (block vertebrae), and 8 (32%) had defects in both the posterior elements and the vertebral bodies (Figure 3D–F). None (0%) of these patients exhibited the small vertebral bodies, enlarged pedicles, or large spinous processes that were seen in all of the patients with FOP (Figure 3A–C).

When compared with wild-type mice (Figure 4A), *Nog* (–/–) mice (Figure 4B, C), which do not express the bone morphogenetic protein (BMP) antagonist noggin, exhibited fusion of the spinous processes and lateral masses but displayed no fusion between the vertebral bodies or between the occiput and C2. The malformations observed in the *Nog* (–/–) mouse are almost identical to the cervical spine radiographic findings in FOP patients.

KF syndrome results from disruption of vertebral segmentation. A related vertebral segmentation disorder, spondylocostal dysostosis/Jarcho-Levin syndrome, is caused by mutations in the *DLL3* gene of the notch pathway,<sup>20</sup> a pathway that plays a key role in mammalian vertebral segmentation. When compared with wild-type mice (Figure 4D), *Dll3*<sup>pm</sup>/*Dll3*<sup>pm</sup> mutant mice (Figure 4E) exhibited malformations of the cervical spine, which are strikingly similar to the segmental defects seen in patients with KF syndrome and identical to those in spondylocostal dysostosis. It should be noted that the cervical spine of the *Nog* (–/–) mouse and the *Dll3*<sup>pm</sup> mutant show very different patterns of fusion, much like the different patterns of fusion seen in the cervical spine of FOP and KF patients (Figure 4).

## ■ Discussion

Fibrodysplasia ossificans progressiva is a disabling developmental disease with progressive heterotopic, endochondral ossification of soft connective tissue and skeletal muscles.<sup>1–11</sup> In addition to the great toe malformations seen in approximately 95% of FOP patients,<sup>1,4,6,7,9–11</sup> joints become ankylosed by heterotopic bone formation, leaving most patients wheelchair-bound by the third decade of life.<sup>1–11</sup> Currently, there is no effective treatment to stop the progression of this genetic disease,<sup>9–11</sup> although a growing body of evidence suggests that interruption of an overactive BMP pathway may be a promising treat-

spinous processes in a patient with FOP (at 2, 8, 10, 14 years of age). White arrowheads indicate progressive fusion. **E–H**, Nonprogressive fusion of the cervical spine in a patient with KF syndrome (at 2, 4, 7, 8 years of age). White arrowheads indicate fusion of spinous processes which show no change in the degree of fusion over time in the patient with KF syndrome.

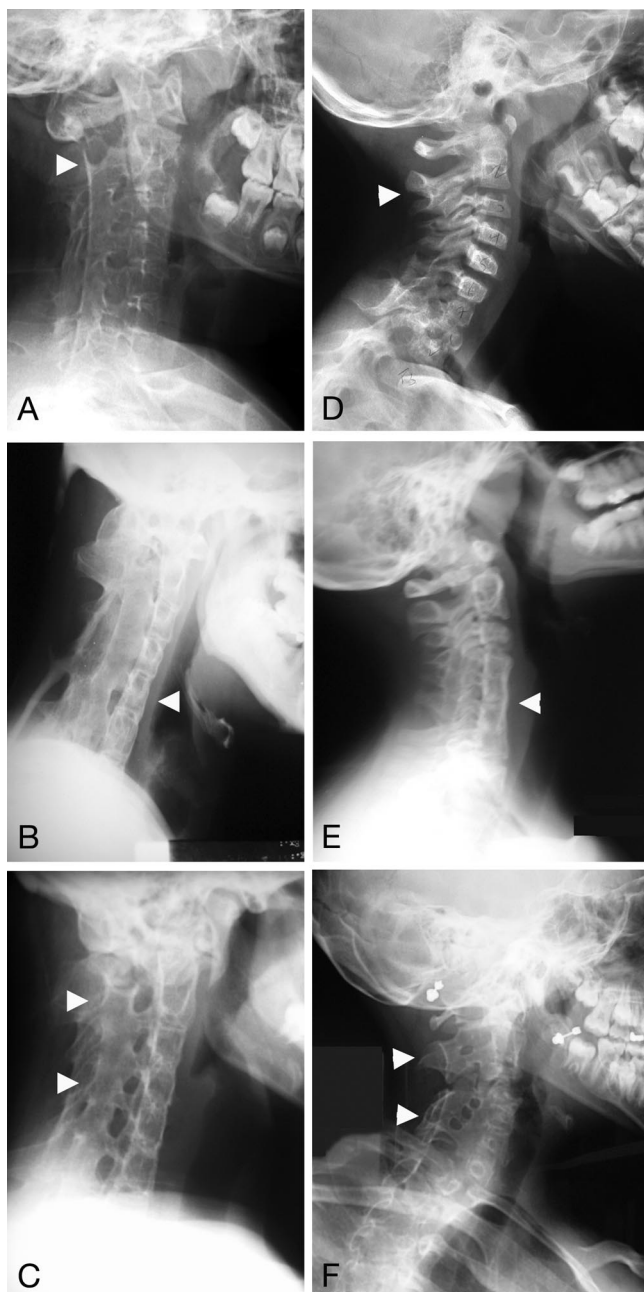


Figure 3. Contrast of the typical orthotopic changes in the cervical spine of three different FOP patients with the typical orthotopic changes in the cervical spine of three different KF patients. **A–C**, The typical fusion pattern seen in the cervical spine of FOP patients. **D–F**, Three different types of defects typically seen in the cervical spines of KF patients. **D**, A single defect; white arrowhead indicates the fusion of C2 to C3. Note that while the FOP patient in **A** also has a fusion of C2 to C3 (white arrow), the fusion is not limited to C2 to C3 but extends to C7. **E**, A block defect indicated with a white arrowhead. Even though at first glance it appears that the FOP patient in **B** also has a block defect, note that the disc spaces are still maintained (white arrowhead) unlike in the block fusion in **E**. **F**, A defect limited mainly to the posterior elements. White arrowheads indicate the most prominent defects, which form an irregular pattern typical of a defect of segmentation. When compared with **C** note that the fusion of the posterior elements in this FOP patient forms a regular pattern (white arrowheads) inconsistent with a defect of segmentation.

ment.<sup>21</sup> Recent studies suggest that the causal mutation of this devastating disease may be found within genes of the BMP signaling pathway, which contribute to embryonic skeletal patterning and postnatal heterotopic ossification.<sup>22–28</sup>

All of the lateral cervical spine radiographs of the 70 FOP patients reviewed in this study showed similar vertebral abnormalities consisting of narrow vertebral bodies, enlarged pedicles, and large spinous processes consistent with previous reports.<sup>2,4,11</sup> In this study, almost all (90%) FOP patients showed some degree of orthotopic or articular fusion of the cervical spine. The remaining 10% were all under the age of 1, suggesting that orthotopic ankylosing is generally not present at birth but progressively develops postnatally. Most likely, FOP patients have cervical spine abnormalities at birth involving abnormal specification of the articular facet joints that progressively fuse as they age. Different patients show various degrees of fusion at different ages, reinforcing variable expression of this disease.<sup>1,9,11</sup>

Most FOP patients present with neck stiffness and decreased range of motion of the cervical spine in the first decade of life,<sup>1,11</sup> often leading to a misdiagnosis of KF. However, a radiographic finding of narrow vertebral bodies, enlarged pedicles, and large spinous processes in the setting of malformed great toes suggests a diagnosis of FOP rather than KF, even before the appearance of soft tissue swellings, which lead to heterotopic ossification. Inappropriate treatment of FOP patients, including invasive procedures, unnecessary biopsies, and intramuscular injections and immunizations,<sup>11</sup> predictably induce increased heterotopic bone formation.<sup>9,11</sup> Therefore, early diagnosis is crucial in preventing undue harm.

Although both FOP and KF patients display malformations of the cervical spine, they clearly do not exhibit the same defects. None of the KF patients reviewed in this study had a cervical spine consisting of narrow vertebral bodies, enlarged pedicles, or large spinous processes. Even the KF patients who had extensive posterior element involvement had patterns of fusion distinct from the FOP patients. The different patterns of fusion seen in these subsets of patients can be traced back to the process of segmentation during development. Patients with KF syndrome exhibited a disorganized pattern of posterior element formation similar to the *Dll3<sup>pu</sup>* mutant mouse that is consistent with a defect in craniocaudal segmentation (Figure 4E). In contrast, the FOP patients exhibited a very uniform pattern of fusion of the posterior elements similar to the *Nog* (–/–) mouse. This uniform pattern of fusion is not suggestive of a defect in spinal segmentation. The ability to recognize the difference between a typical FOP cervical spine (Figure 3A–C) and a typical KF cervical spine (Figure 3D–F) will help lead the clinician to an appropriate diagnosis and ensure that the patient receives appropriate treatment.

In the past two decades, research in developmental biology has uncovered molecular pathways that control developmental patterning. For example, the secreted



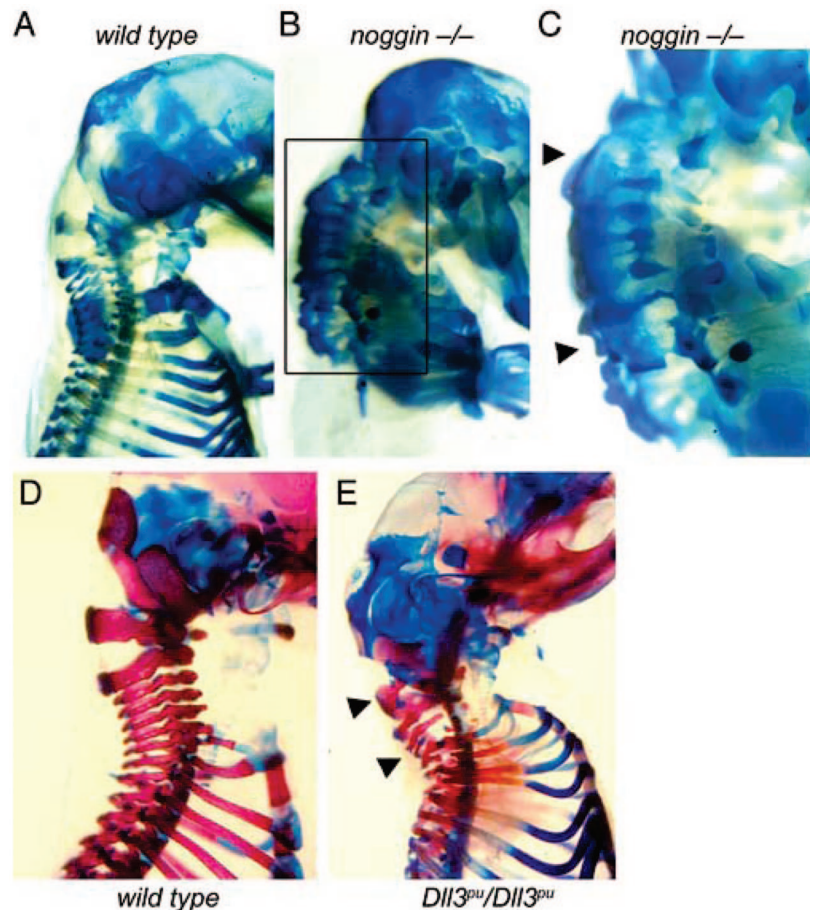


Figure 4. Cervical vertebral abnormalities in mouse homozygous deletion *Nog* ( $-/-$ ) mutants compared with  $\delta$ -like 3 (*Dll3*<sup>pu</sup>/*Dll3*<sup>pu</sup>) homozygous mutants. **A–C**, Comparison of wild-type (**A**) and *Nog* ( $-/-$ ) neonatal mice skeletons (**B** and **C**). Black arrowheads indicate areas of fusion. Note the similarities (posterior element fusion with no vertebral body fusion and lack of fusion between occiput and C2) between the *Nog* ( $-/-$ ) mouse and FOP patients. **D** and **E**, Comparison of wild-type (**D**) with *Dll3*<sup>pu</sup> mutant skeleton (**E**). Black arrowheads indicate defects in segmentation, similar to those seen in KF patients. Cartilage (blue) is stained with Alcian Blue (**A–E**) and bone (red) is stained with Alizarin Red (**D** and **E**).

morphogen sonic hedgehog (SHH) is released from the developing notochord and plays a key role in ventralizing the embryonic spinal cord and somites, which are the precursors to the vertebrae.<sup>29</sup> Loss of the *Shh* gene in mice leads to lack of the entire vertebral column, illustrating the key role of this factor in spinal development.<sup>30</sup> Interestingly, in the mouse, disruption of both the *Zic1* and *Gli3* genes, which function downstream of *Shh*, leads to vertebral fusions of the posterior elements farthest from the source of sonic hedgehog factor.<sup>31</sup>

Genes in the BMP pathway also appear to play a major role in developmental patterning and spinal development. The BMP pathway encompasses a large number of genes, including ligands (BMP/TGF- $\beta$  secreted factors, activins), receptors (BMP receptors, TGF- $\beta$  receptors, activin receptors), stimulatory and inhibitory transcription factors (SMAD proteins), and secreted antagonists (noggin, chordin, gremlin, cerberus, follistatin).<sup>22,29</sup> Noggin encodes a potent BMP antagonist that acts by binding to BMP and preventing it from binding to its transmembrane receptor.<sup>16,32–34</sup> Disruption of noggin in the mouse results in numerous skeletal defects resulting from excess cartilage formation. Most notably, *Nog* knockout mice fail to appropriately specify diarthrodial joints and exhibit extensive cervical vertebral fusions, particularly C2 and below, nearly identical to the defects described here in FOP patients<sup>16</sup> (Figure 4). Al-

though *NOG* mutations occur in patients with symphalangism and syndactyly,<sup>35</sup> as well as tarsal-carpal coalition syndrome<sup>36</sup> and stapes ankylosis with broad thumb and toes,<sup>37</sup> *noggin* mutations have not been confirmed in FOP patients.<sup>26</sup> It is evident, however, that although FOP is not caused by *NOG* mutations, enhanced activation of the BMP pathway is pathognomonic of FOP and leads to phenotypic findings in FOP patients that are similar (although less severe) than those seen in the *Nog* knockout mouse. Most notably, the pattern of orthotopic fusion of the cervical spine in the mouse model is nearly identical to that seen in FOP patients (Figure 4).<sup>2,4,6,7,9–11,16</sup>

Thus, while *NOG* is not disrupted in FOP, the vertebral defects observed in FOP patients are highly consistent with embryonic dorsoventral shifts resulting from overactivity of BMP signaling. Interestingly, BMP pathway genes are involved in patterning of the limbs,<sup>22,29</sup> which is consistent with the congenital great toe malformations found in FOP patients.<sup>38</sup> Discovery of the mutated gene for FOP and identification of the developmental pathways leading to its pathology will ultimately explain the congenital cervical vertebral defects and great toe malformations observed in FOP patients and lead to improved treatment strategies.

Heterotopic ossification of the cervical spine in FOP patients, as shown in Figure 1, has clearly been described



in previous reports.<sup>1,2,4,6-11</sup> This paper differentiates between these defects and the altered pattern of orthotopic ossification seen in the cervical spine of patients with FOP, which may result from different functions of the FOP gene between regulation of ossification postnatally and embryonic patterning of the skeleton. Identification of the FOP gene could have implications in the treatment of patients who require therapeutic spine fusions, building on recent studies demonstrating the successful use of BMP pathway gene products to induce spinal fusion.<sup>39</sup> Better understanding of the BMP pathway may provide insight into the molecular genetics of both orthotopic and heterotopic ossification and will likely provide further insight into how to use these gene products therapeutically to augment patients who need spinal fusions.

## ■ Conclusion

The ability to distinguish between the appearance of an FOP cervical spine and a Klippel-Feil cervical spine is imperative for early diagnosis and appropriate treatment for these patient groups. Understanding the possible etiology of these abnormalities in terms of developmental pathways is important to further differentiate cervical spines of FOP and KF patients, to help elucidate the genetic cause of fibrodysplasia ossificans progressiva, and to develop appropriate treatment strategies for patients with FOP.

## ■ Key Points

- In a radiographic analysis of 70 fibrodysplasia ossificans progressiva patients and 33 Klippel-Feil syndrome patients, congenital cervical spine abnormalities were observed but were distinctly different in each group.
- The anomalies seen in the cervical spine of the fibrodysplasia ossificans progressiva patient group were strikingly similar to those seen in mice with homozygous deletions of the gene encoding noggin, a potent bone morphogenetic protein antagonist.
- Understanding the etiology of these observed abnormalities will help to elucidate the genetic cause of fibrodysplasia ossificans progressiva.

## Acknowledgments

The authors thank the families for participating in the Congenital Vertebral Malformations and Fibrodysplasia Ossificans Progressiva studies; Drs. Denis Drummond, John Flynn, Malcolm Ecker, and Richard Davidson for contributing cases for this study; and Julia Lou, Stephan Pill, Akshay Mehta, Yana Tsygansky, Mizuho Mimoto, and Vanessa Garcia for technical assistance.

## References

- Cohen RB, Hahn GV, Tabas JA, et al. The natural history of heterotopic ossification in patients who have fibrodysplasia ossificans progressiva: a study of forty-four patients. *J Bone Joint Surg Am* 1993;75:215-9.
- Connor JM, Smith R. The cervical spine in fibrodysplasia ossificans progressiva. *Br J Radiol* 1982;55:492-6.
- Corfield L, Hampton R, McCullough CJ. Wrist arthrodesis following ulnar bar excision in fibrodysplasia ossificans progressiva. *J Hand Surg Br* 2000;25:223-4.
- Cremin B, Connor JM, Beighton P. The radiological spectrum of fibrodysplasia ossificans progressiva. *Clin Radiol* 1982;33:499-508.
- Kaplan FS, Gannon FH, Hahn GV, et al. Pseudomalignant heterotopic ossification. *Clin Orthop* 1998;346:134-40.
- Kaplan FS, Strear CM, Zasloff MA. Radiographic and scintigraphic features of modeling and remodeling in the heterotopic skeleton of patients who have fibrodysplasia ossificans progressiva. *Clin Orthop* 1994;304:238-47.
- Mahboubi S, Glaser DL, Shore EM, et al. Fibrodysplasia ossificans progressiva. *Pediatr Radiol* 2001;31:307-14.
- Shah PB, Zasloff MA, Drummond D, et al. Spinal deformity in patients who have fibrodysplasia ossificans progressiva. *J Bone Joint Surg Am* 1994;76:1442-50.
- Smith R. Fibrodysplasia (myositis) ossificans progressiva: clinical lessons from a rare disease. *Clin Orthop* 1998;346:7-14.
- Smith R, Athanasou NA, Vipond SE. Fibrodysplasia (myositis) ossificans progressiva: clinicopathological features and natural history. *QJM* 1996;89:445-6.
- Kaplan FS, Shore EM, Connor JM. Fibrodysplasia ossificans progressiva. In: Royce PM, Steinmann B, eds. *Connective Tissue and Its Heritable Disorders: Molecular, Genetic, and Medical Aspects*, 2nd ed. New York: Wiley-Liss, 2002:827-40.
- Baba H, Maezawa Y, Furusawa N, et al. The cervical spine in the Klippel-Feil syndrome: a report of 57 cases. *Int Orthop* 1995;19:204-8.
- Falliner A, Drescher W, Brossmann J. The spine in fibrodysplasia ossificans progressiva: a case report. *Spine* 2003;28:E519-22.
- Clarke RA, Catalan G, Diwan AD, et al. Heterogeneity in Klippel-Feil syndrome: a new classification. *Pediatr Radiol* 1998;28:967-74.
- Thomsen MN, Schneider U, Weber M, et al. Scoliosis and congenital anomalies associated with Klippel-Feil syndrome types I-III. *Spine* 1997;22:396-401.
- Brunet LJ, McMahon JA, McMahon AP, et al. Noggin, cartilage morphogenesis, and joint formation in the mammalian skeleton. *Science* 1998;280:1455-7.
- McLeod MJ. Differential staining of cartilage and bone in whole mouse fetuses by alcian blue and alizarin red S. *Teratology* 1980;22:299-301.
- Gruneberg H. Genetical studies on the skeleton of the mouse: XXIX. *Pudgy. Genet Res Camb* 1961;2:384-93.
- Kusumi K, Sun ES, Kerrebrock AW, et al. The mouse pudgy mutation disrupts Delta homologue Dll3 and initiation of early somite boundaries. *Nat Genet* 1998;19:274-8.
- Bulman MP, Kusumi K, Frayling TM, et al. Mutations in the human delta homologue, DLL3, cause axial skeletal defects in spondylocostal dysostosis. *Nat Genet* 2000;24:438-41.
- Glaser DL, Economides AN, Wang L, et al. In vivo somatic cell gene transfer of an engineered Noggin mutein prevents BMP4-induced heterotopic ossification. *J Bone Joint Surg Am* 2003;85:2332-42.
- Kaplan FS, Ahn J, Serrano de la Peña LS, et al. Fibrodysplasia ossificans progressiva: deciphering the molecular pathways of ectopic skeletogenesis. In: Rosier RN, Evans CH, eds. *Molecular Biology in Orthopaedics*. Rosemont, IL: American Academy of Orthopaedic Surgeons, 2001:95-105.
- Shafritz AB, Shore EM, Gannon FH, et al. Overexpression of an osteogenic morphogen in fibrodysplasia ossificans progressiva. *N Engl J Med* 1996;335:555-61.
- Olmsted EA, Kaplan FS, Shore EM. Bone morphogenetic protein-4 regulation in fibrodysplasia ossificans progressiva. *Clin Orthop* 2003;408:331-43.
- Gannon FH, Kaplan FS, Olmsted E, et al. Bone morphogenetic protein 2/4 in early fibromatous lesions of fibrodysplasia ossificans progressiva. *Hum Pathol* 1997;28:339-43.
- Xu MQ, Feldman G, Le Merrer M, et al. Linkage exclusion and mutational analysis of the noggin gene in patients with fibrodysplasia ossificans progressiva (FOP). *Clin Genet* 2000;58:291-8.
- Gong Y, Krakow D, Marcelino J, et al. Heterozygous mutations in the gene encoding noggin affect human joint morphogenesis. *Nat Genet* 1999;21:302-4.
- Ahn J, Serrano de la Peña L, Shore EM, et al. Paresis of a bone morphogenetic protein-antagonist response in a genetic disorder of heterotopic skeletogenesis. *J Bone Joint Surg Am* 2003;85:667-74.
- Kornak U, Mundlos S. Genetic disorders of the skeleton: a developmental approach. *Am J Hum Genet* 2003;73:447-74.
- Chiang C, Litingtung Y, Lee E, et al. Cyclopia and defective axial patterning in mice lacking Sonic hedgehog gene function. *Nature* 1996;383:407-13.

31. Aruga J, Mizugishi K, Koseki H, et al. Zic1 regulates the patterning of vertebral arches in cooperation with Gli3. *Mech Dev* 1999;89:141–50.
32. Zimmerman LB, Harland RM. The Spemann organizer signal noggin binds and inactivates bone morphogenetic protein 4. *Cell* 1996;86:599–606.
33. Groppe JGJ, Wiater E, Rodriguez-Leon J, et al. Structural basis of BMP signalling inhibition by the cystine knot protein Noggin. *Nature* 2002;420:636–42.
34. Groppe JGJ, Wiater E, Rodriguez-Leon J, et al. Structural basis of BMP signaling inhibition by Noggin, a novel twelve-membered cystine knot protein. *J Bone Joint Surg Am* 2003;85:52–8.
35. Marcelino J, Sciortino CM, Romero MF, et al. Human disease-causing NOG missense mutations: effects on noggin secretion, dimer formation, and bone morphogenetic protein binding. *Proc Natl Acad Sci USA* 2001;98:11353–8.
36. Dixon ME, Armstrong P, Stevens DB, et al. Identical mutations in NOG can cause either tarsal/carpal coalition syndrome or proximal symphalangism. *Genet Med* 2001;3:349–53.
37. Brown DJ, Kim TB, Petty EM, et al. Autosomal dominant stapes ankylosis with broad thumbs and toes, hyperopia, and skeletal anomalies is caused by heterozygous nonsense and frameshift mutations in NOG, the gene encoding noggin. *Am J Hum Genet* 2002;71:618–24.
38. Dunn NR, Hargett LK, Schrick JJ, et al. Haploinsufficient phenotypes in Bmp4 heterozygous null mice and modification by mutations in Gli3 and Alx4. *Dev Biol* 1997;188:235–47.
39. Boden SD, Kang J, Sandhu H, et al. Use of recombinant human bone morphogenetic protein-2 to achieve posterolateral lumbar spine fusion in humans: a prospective, randomized clinical pilot trial. 2002 Volvo Award in clinical studies. *Spine* 2002;27:2662–73.

## Strut Allograft Union and Remodeling Using rhBMP-2 in a Spinal Corpectomy Model

Klane K. White, MD,\* Maneesh Bawa, MD,\* Jae-Sung Ahn, MD,†  
Christine L. Farnsworth, MS,‡ Frances D. Faro, MD,\* Andrew T. Mahar, MS,‡  
Michelle Wedemeyer, BS,‡ Steven R. Garfin, MD,\* and Peter O. Newton, MD‡

**Study Design.** Growth factor in a collagen sponge carrier was compared to autograft, both packed within an allograft strut following corpectomy in a bovine spinal model.

**Objective.** To evaluate incorporation of cortical strut allografts after lumbar corpectomy, comparing augmentation with recombinant human bone morphogenetic protein-2 (rhBMP-2) to local autograft.

**Summary of Background Data.** Autogenous strut grafts are the gold standard for successful fusion in reconstruction following corpectomy; however, significant donor site morbidity can occur. Recent studies describing consistently successful anterior interbody fusions with BMP augmentation suggest an exciting prospect of unlimited and potent grafting material for these difficult fusions.

**Methods.** Sixteen calves underwent L3 corpectomy with instrumented strut allograft reconstruction. The rhBMP-2 impregnated collagen sponges filled the empty medullary canal of the allograft in 8 animals. Eight animals had the allograft strut filled with local autogenous bone. After 4 months, the lumbar spines were harvested for radiographic, biomechanical, and histologic evaluation.

**Results.** Computerized tomography revealed allograft fusion in 7 of 8 autograft specimens and 8 of 8 BMP specimens. The BMP treated group had denser bone at the ends of the cortical allograft, but a central void persisted. Autograft filled struts maintained a more uniform distribution of less organized bone throughout the strut canal. Histologic assessment verified remodeling and incorporation of the allografts for both groups. Biomechanical testing confirmed no significant difference in fusion strength between groups.

**Conclusions.** Large cortical strut allografts (after lumbar corpectomy) supplemented with rhBMP-2 had incor-

poration and fusion strength comparable to allografts enhanced with cancellous autograft.

**Key words:** lumbar spine, corpectomy, bone morphogenetic proteins, allograft, autograft. **Spine** 2005;30:1386–1395

Anterior corpectomy with instrumentation and fusion is common in the treatment of fractures, tumors, infections, and kyphotic deformities of the thoracic and lumbar spine.<sup>1–3</sup> Presently, autogenous strut grafts from the iliac crest, rib, or fibula are the gold standard for achieving successful fusion of this construct. However, each of these sources of bone can result in significant donor site morbidity.<sup>4–7</sup> In some situations, there may not be enough autogenous bone available for reconstruction. For these reasons, allogeneous sources of bone graft have been used with increasing frequency. Allograft bone has become widely available, and provides the surgeon many options regarding graft size and shape.<sup>5,8</sup>

Clinical and experimental studies assessing allograft and autograft sources of bone have shown mixed results when comparing their use in spinal fusion; however, it is clear that autogenous bone generally provides higher rates of fusion.<sup>9–12</sup> The relative superiority of autogenous bone sources is thought to be a result of the presence of both osteogenic cells and endogenous growth factors. Bone morphogenetic proteins (BMP) are important in osteoinduction, initiation of osteogenic cellular differentiation, and proliferation. The use of recombinant human (rh) BMP-2 with a collagen sponge carrier has been successful clinically for achieving anterior lumbar interbody arthrodesis when combined with allograft cortical bone.<sup>13–18</sup>

BMP augmentation of larger allografts has also been successful for the treatment of long bone segmental defects and tibial nonunions.<sup>19–22</sup> Although studied extensively in long bone reconstruction, large cortical allograft incorporation and remodeling have received less attention for spinal applications. Long structural allografts have been used successfully to reconstruct anterior spinal column defects.<sup>5,8,23,24</sup> In these studies, radiographic appearances suggested that resorption and remodeling occurred at the ends of the graft. However, standard radiographs are not reliable for evaluating the viability of these grafts. It is likely that these large grafts in the spine, similar to large allografts elsewhere in the body, are not fully incorporated, leaving a large portion of the graft as a necrotic center.<sup>25,26</sup>

From the Department of Orthopaedic Surgery, \*University of California San Diego, San Diego, CA, †Chungnam National University, Republic of Korea, and ‡Children's Hospital & Health Center, San Diego, CA.

Acknowledgment date: May 12, 2004. First revision date: July 29, 2004. Acceptance date: August 2, 2004.

Supported by Medtronic Sofamor Danek.

The device(s)/drug(s) that is/are the subject of this manuscript is/are not FDA-approved for this indication and is/are not commercially available in the United States.

Corporate/Industry funds were received in support of this work. Although one or more of the author(s) has/have received or will receive benefits for personal or professional use from a commercial party related directly or indirectly to the subject of this manuscript, benefits will be directed solely to a research fund, foundation, educational institution, or other nonprofit organization which the author(s) has/have been associated.

Address correspondence and reprint requests to Peter O. Newton, MD, 3030 Children's Way, Suite 410, San Diego, CA 92123-4293; E-mail: pnewton@chsd.org



The purpose of this study was to evaluate the value of rhBMP-2, compared to cancellous autogenous bone, for augmenting union and remodeling of a cortical allograft used to reconstruct an L3 corpectomy defect in a bovine model. The null hypothesis was that there was no radiographic, histologic, or biomechanical differences between strut allografts augmented with rhBMP-2 and those augmented with local autograft following a 4-month incorporation period.

## Materials and Methods

**Overview.** This study was approved by the Animal Subjects Committee of the authors' institution and complied with all regulations for the humane treatment of animals. An immature bovine model was chosen because the size of the lumbar spine is similar to the adult human lumbar spine, allowing for use of standard implants and requiring a large allograft strut for corpectomy reconstruction. One-month-old male calves were acquired and housed at a large animal facility for one week before the procedure. A total of 16 calves underwent L3 corpectomy and strut graft reconstruction. The rhBMP-2 impregnated absorbable collagen sponges were placed in the strut allograft medullary canal of 8 animals. Eight "control" animals had the allograft filled with autogenous cancellous bone from the excised vertebral body. The procedures were alternated to account for progressive surgical experience.

### Surgical Protocol

The calves were not fed for 12 hours before the procedure and were given 0.01 mg/kg buprenorphine intramuscularly for pain treatment and 1 g cefazolin intramuscularly before surgery. Calves were sedated with 0.3 mg/kg diazepam intravenously; induction with 2 mg/kg intravenous propofol (Diprivan®, AstraZeneca Pharmaceuticals LP, Wilmington, DE) was followed by intubation with an 8–10 mm inner diameter, cuffed endotracheal tube. Anesthesia was maintained with 1% to 3% volatilized isoflurane until the end of the procedure. The subjects were given a maintenance infusion of lactated ringers solution at a rate of 10 cc/kg/h for the first hour, then 5 cc/kg/h until the conclusion of surgery. In a lateral position, each calf's right flank was prepared with an iodine-based scrub, and the surgical site was appropriately draped. An anterior retroperitoneal approach to the lumbar spine was used through a curvilinear skin incision. Once adequate exposure of L2–L4 was obtained, the L2–L4 segmental vessels were ligated, and a corpectomy of L3 was performed.

Fresh frozen tibial allografts, previously harvested from another group of same-age calves, were thawed at room temperature in normal saline with cefazolin (average length  $36.5 \pm 1.5$  mm; range 33–38). The allograft struts had been harvested with sterile technique, removing all marrow elements and periosteum. The grafts were then wrapped in saline soaked gauze, double bagged, and stored at  $-20^{\circ}\text{C}$ . In the control group, the hollow medullary canal of the cortical allograft was packed with local cancellous autograft from the excised L3 vertebra. In the treatment group, 5 rhBMP-2 impregnated  $1 \times 2$ -in collagen sponges were placed in the cortical allograft (1.4 mL of 0.43 mg/mL rhBMP-2 per sponge or 0.602 mg BMP per sponge) (InFuse®, Medtronic Sofamor Danek, Memphis, TN; Helistat® absorbable collagen sponge, Integra Life Sciences, Plainsboro, NJ). The total dose of BMP was 3.0 mg per graft. The strut graft reconstruction was then stabilized with an an-

terior ATL Z-plate (Medtronic Sofamor Danek) according to the manufacturer's guidelines. The wound was closed using absorbable suture. By veterinarian recommendation, the calves were castrated at the time of the spinal procedure to reduce herd aggression as the calves matured.

Postoperative analgesia was provided with intramuscular injections of 0.01 mg/kg buprenorphine delivered every 12 hours for the first 3 days. Postoperative antibiotics consisted of 500 mg cefazolin intramuscularly twice a day for 3 days. Animals were housed in climate controlled isolation pens for one week following surgery and then returned to outdoor pens with ad lib activity. Following a 4-month survival period, allowing fusion to occur according to previous studies with BMP augmented spine fusion in nonprimate animals,<sup>13,16,27–30</sup> the animals were sedated (30 mg xylazine intramuscularly), and then euthanized by administration of 0.3 mL/kg Beuthanasia D Special® intravenously (Schering-Plow Animal Health Corp., Kenilworth, NJ). The lumbar spines were harvested en bloc from L2–L4. The first 2 spines from each surgical group were stored in 10% buffered formalin before histologic processing. Six spines from each surgical group (BMP and autograft) were then stored frozen at  $-20^{\circ}\text{C}$  before biomechanical testing.

### Radiographic Assessment

Posteroanterior and lateral radiographs were taken immediately after surgery. Following harvest, a second set of radiographs was obtained of each spine. All instrumentation was removed before performing spiral computerized tomography (CT), which were obtained with 1-mm slices, and reformatted for coronal and sagittal reconstructions. Two reviewers who were blinded to the treatment group evaluated all studies for the presence of fusion.

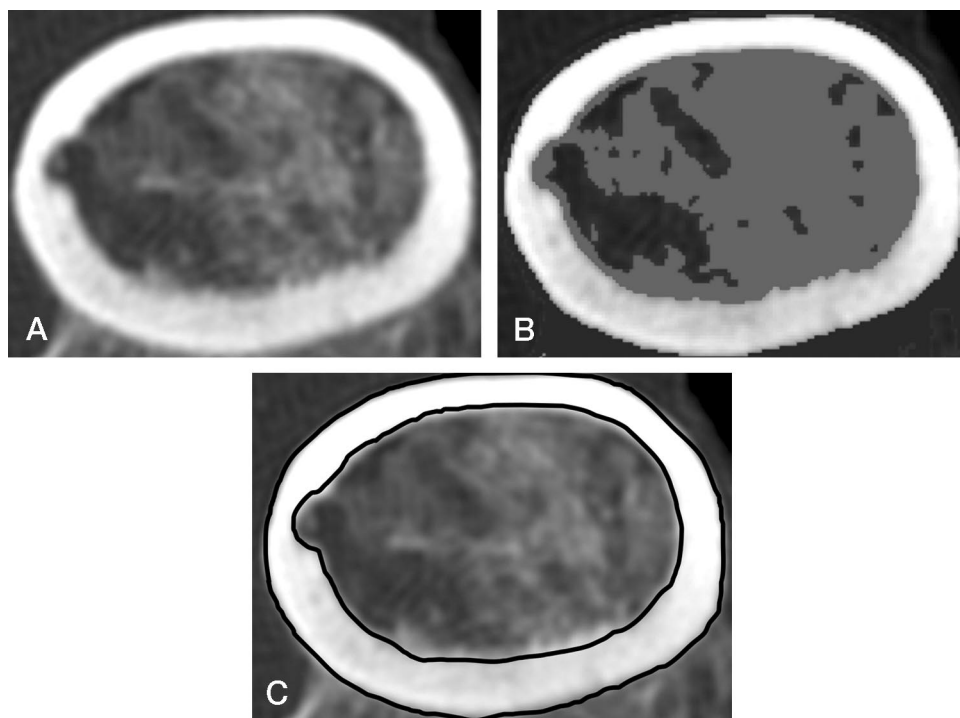
Digital transverse CT images also underwent image analysis (Scion Image for Windows, version 4.0, Scion Corp., Frederick, MD) to determine the amount of bone within the strut allograft medullary canal and the degree of cortex resorption (Figure 1). Gray-scale thresholds for normal vertebral cancellous bone were measured and applied digitally to 7 sequential CT slices from the superior edge to the inferior edge of the allograft to determine the amount of "bone" within each graft. Each pixel within the inner cortex of the strut graft that met or exceeded the threshold density was digitally selected. The selected area was divided by the total cross sectional area of the allograft canal to determine the percentage of the area that met the threshold density for bone ("percent area of bone" value). The average percent area of bone at each of the 7 levels along the length of the allograft was determined and compared between treatments with a repeated measures analysis of variance (ANOVA) with a significance of  $\alpha = 0.01$ .

The amount of strut allograft cortex resorption was also measured on the same 7 CT images for each specimen. The original preincorporation cortex boundaries were approximated, and the average gray-scale value for the area of cortex only (excluding medullary canal) was calculated. A scale ranging from 0 to 240 was used for these density measures, with 0 representing the highest density (white on CT) and 240 the lowest density (black). A single technician who was blinded to the treatment graded all digitized CT images. Comparisons between treatment groups were made by repeated measures ANOVA with a statistical significance of  $\alpha = 0.01$ .

### Qualitative Histology

The first 2 specimens in each surgical group were predesignated for histologic analysis. Each of the 4 specimens, including the

Figure 1. Digital image analysis methodology. **A**, Transverse CT image through a strut graft 4 months after placement. **B**, The shaded area has a gray-scale value more than or equal to adjacent cancellous bone, and was digitally selected and measured to show the presence of cancellous bone within the graft. **C**, Transverse CT image of the same spine where the outer and inner (solid black line) edges of the cortex have been digitally selected. Average gray-scale value was measured for the area within these boundaries to determine the amount of resorbed cortical bone.



distal half of L2 to the proximal half of L4, were sectioned into right and left halves along the midsagittal plane. The left half of all 4 spines were dehydrated in ascending concentrations of ethyl alcohol, cleared in xylene, and embedded in methylmethacrylate, using previously published protocols.<sup>31,32</sup> After embedding in methylmethacrylate, the spines were cut in the sagittal plane with a high speed, water cooled, low feed bone saw with a diamond impregnated blade. Contact radiographs were made of each section with high-resolution film. The midsagittal sections were ground and polished to a thickness of 50–100  $\mu\text{m}$ .

The right half of all 4 spines were decalcified in phosphate buffered neutral ethylenediaminetetraacetic acid. Paraffin-embedded blocks were made from the midsagittal cut, with 3 each at the proximal and distal graft-host bone interfaces (anterior, mid, posterior), and 3 blocks from the mid graft region. Each was sectioned at a thickness of 5  $\mu\text{m}$ , mounted on glass slides, and stained with hematoxylin and eosin, and Sanderson trichrome stains. The histologic evaluation by 2 evaluators who were blinded to the treatment group was qualitative and made primarily to corroborate the findings measured quantitatively on the CT. No attempt was made to measure statistically any histologic differences between the groups given the limited number of histologic specimens.

### Biomechanical Analysis

Twelve lumbar spines (6 from each surgical group) underwent biomechanical testing of the fused segments. The ends of the L2 and L4 vertebrae were fixed in a 2-part epoxy resin within custom designed fixation jigs. The specimens were mounted in an MTS 858 Mini-bionix biaxial testing machine (MTS Systems, Minneapolis, MN) for biomechanical evaluation. Alignment in the testing machine was such that the mechanical axis of testing was centered within the canal of the cephalad and caudal vertebral bodies. Nondestructive testing in flexion (5 Nm), extension (5 Nm), lateral bending (5 Nm), and torsion (5 Nm with 100 N axial load) was performed. Five loading cycles

were completed (3 cycles of preload followed by 2 cycles for data acquisition). The torque and angular displacement data were collected at 10 Hz for the duration of each test. Range of motion and stiffness were measured for the cyclic tests. Ultimate torque to failure was then performed at an angular rate of 0.5° per second. Data were analyzed using 1-way ANOVA. Statistical significance was  $\alpha = 0.01$ .

## Results

### Radiographic Assessment

Plain films performed immediately after surgery revealed well-fixed, intact instrumentation; all strut grafts appeared well placed. After harvest, evaluation of CT revealed allograft fusion with the vertebral endplates in all BMP treated specimens. However, if 2 of the BMP treated calves that survived only 12 weeks (one with an obvious pseudarthrosis) are included, 9 of 10 animals treated with BMP had successful graft fusion. In the autograft treated group, 7 of the 8 spines had allograft fusion at the endplates (Figure 2). The one pseudarthrosis in the autograft group resulted in a kyphotic deformity, with the entire allograft encased in bone, and no continuity between the graft and the vertebral endplates or surrounding bone. No bone overgrowth was seen in any of the other animals in the autograft or BMP treated groups.

Image analysis of CT showed a trend of different patterns of radiopacity within the strut allografts between treatment groups ( $P = 0.05$ ), although the actual values of radiopacity were not significantly different. As shown in Figure 3, the BMP treated group had a higher percent area of bone at the ends of the allograft compared to the middle of the graft, while the autograft filled struts maintained a uniform distribution of bone throughout the

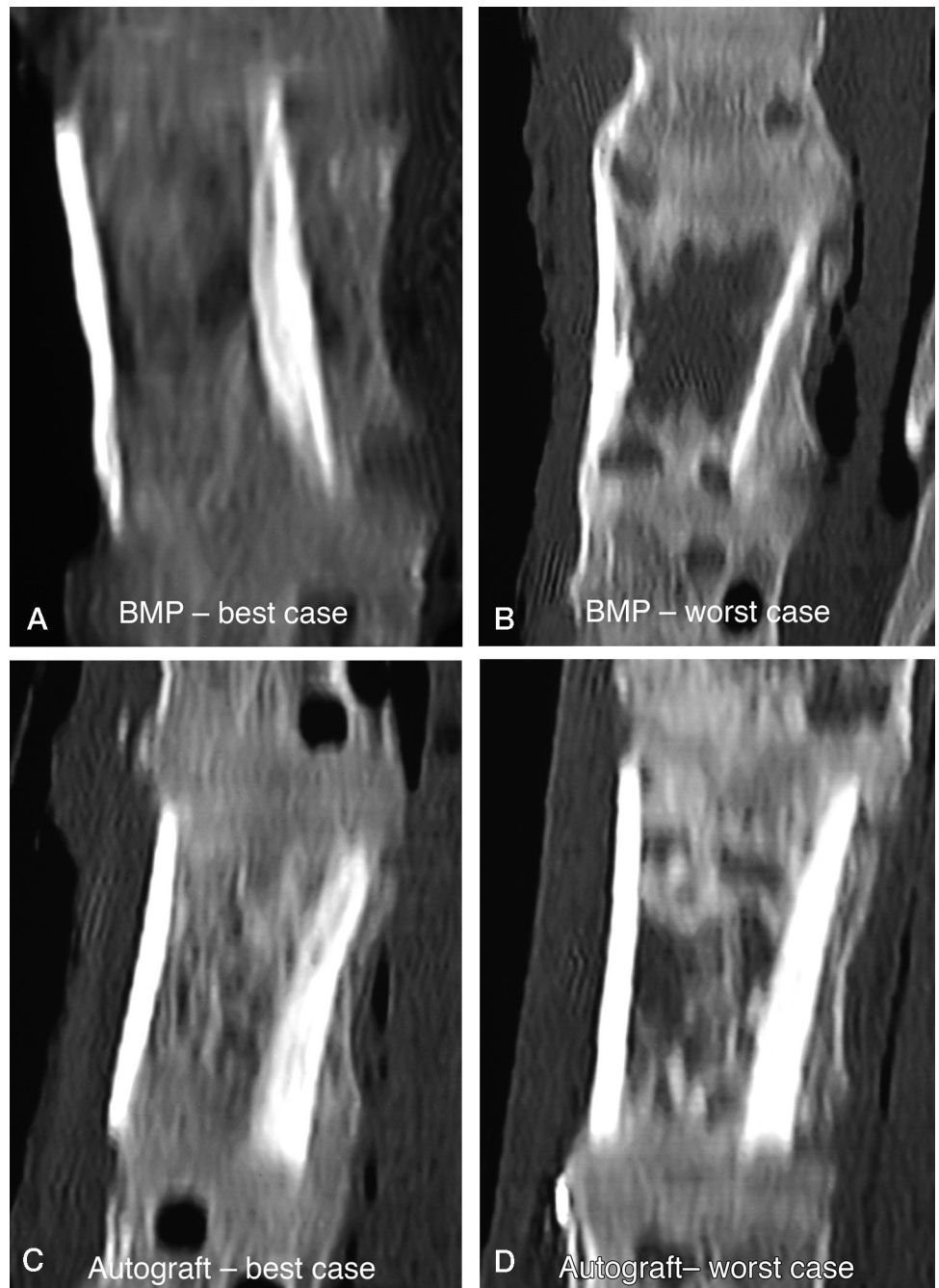


Figure 2. Midsagittal CT images of lumbar corpectomy site 4 months after surgery. Parts **A** and **B** are of grafts filled with *rhBMP-2* impregnated sponges. **A**, Best case. **B**, The case of least amount of bone formation within the strut allograft. Parts **C** and **D** are of autograft filled grafts. **C**, The “best” case. **D**, The “worst” case.

length of the struts. In both groups, there was evidence of resorption of the cortical strut graft adjacent to the endplates (Figure 4), with a larger extent of resorption on the ends of the graft than in the center. There was no difference in cortical resorption between the 2 treatment groups ( $P = 0.99$ ).

#### Histologic Assessment

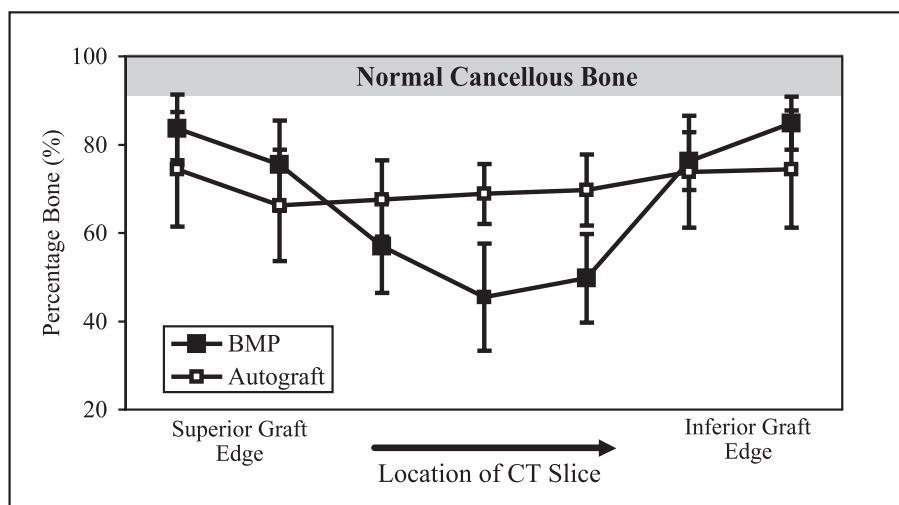
Three spines were evaluated histologically, including 2 BMP treated spines and 1 autograft treated spine. The second spine from the autograft group slated for histologic analysis had the pseudarthrosis found on CT analysis. On discovery of the pseudarthrosis, the processing of this specimen was halted.

Contact radiographs of the nondecalcified sections provided a gross view of bone formation into the medullary canal of the allograft and fusion at the endplates adjacent to the allograft in both groups. These findings matched the patterns of bone growth and cortex resorption found in the quantitative CT analysis of the specimens (Figure 5, available for viewing online through ArticlePlus only).

Hematoxylin and eosin stained slides of the allograft-host bone interface showed similar incorporation of the allograft in both groups. The lacunae were empty within the cortex of the allograft bone of both the BMP and autograft treated specimens. The periphery of the allo-



Figure 3. Percent area of radiopacity measured from sequential transverse CT slices for the grafts filled with BMP and for those with autograft (values shown are mean  $\pm$  standard error;  $P = 0.05$ ). It is not possible from this analysis to distinguish "old" or "new" bone formation in the autograft group. In the BMP group, all bone is presumed "new" bone because only BMP soaked sponges were placed into the strut grafts at surgery. The range of normal bone (shown by the gray bar) is the percent of radiopacity within vertebral bodies adjacent to the graft ( $97.5\% \pm 4\%$ ).



graft was surrounded by osteoblasts and osteoclasts with scalloping of the cortical bone (Howship lacunae), showing allograft resorption. In addition, neovascularization, canals of new bone ingrowth with occupied lacunae and intervening cement lines, confirmed remodeling and incorporation of the allograft cortex.

Despite the limited number of histologic samples, the findings within the medullary canal of the allograft seemed to differ between treatment groups. In all 3 specimens, there was viable appearing trabecular bone. In the BMP treated samples, this trabecular bone (Figure 6A) represented new bone formation because the canal of the allograft was only filled with BMP impregnated collagen sponges at the index procedure. In these BMP specimens, a small central region remained largely devoid of bone (Figure 6B). Small spicules of new bone surrounded this void, with the trabeculae appearing thicker and denser the farther from this zone the specimen was analyzed. Within the central "void," a homogenous, largely acellular matrix with a few penetrating vessels remained that was presumed to be the remnants of the collagen sponge. Vascular channels were identified at the interface between the central void and the newly formed trabecular bone, and penetrating into the central void. This pattern

of a central void and increasing concentration of trabecular bone approaching the host-allograft interface paralleled the pattern seen in the CT analysis. The medullary canal of the autograft filled strut had viable bone throughout the length of the allograft; however, spicules of necrotic bone with empty lacunae were present centrally. The necrotic bone was surrounded by viable new bone with active remodeling of the cancellous autograft used to fill the strut.

#### Biomechanical Assessment

Biomechanical testing under physiologic loads showed similar stability between the BMP treated and the autograft groups (Figure 7). The nondestructive testing of bending and torsional stiffness for the BMP and autograft groups for each test were (mean  $\pm$  standard deviation): flexion ( $4.6 \pm 1.8$  vs.  $3.2 \pm 1.4$  Nm/deg,  $P = 0.16$ ); extension ( $5.3 \pm 2.4$  vs.  $3.1 \pm 1.4$  Nm/deg,  $P = 0.09$ ); lateral bending ( $4.8 \pm 1.0$  vs.  $3.4 \pm 2.0$  Nm/deg,  $P = 0.16$ ); and torsion ( $7.2 \pm 1.7$  vs.  $5.1 \pm 2.0$  Nm/deg,  $P = 0.14$ , respectively). The lower 95% confidence interval for the BMP group was higher than mean value for the autograft treated animals for torsional stiffness and lateral bending stiffness. No statistical difference be-

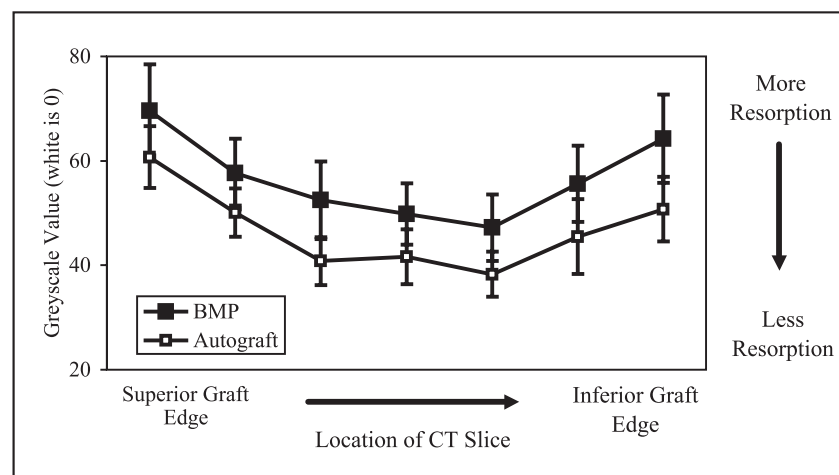


Figure 4. Increased resorption is represented by higher gray-scale values (mean  $\pm$  standard error). Although the BMP group appeared to have more resorption at all levels than the autograft group, this finding was not supported statistically ( $P = 0.71$ ,  $1-\beta = 0.07$ ). Both groups show higher resorption of the strut graft adjacent to the endplates than in the center.

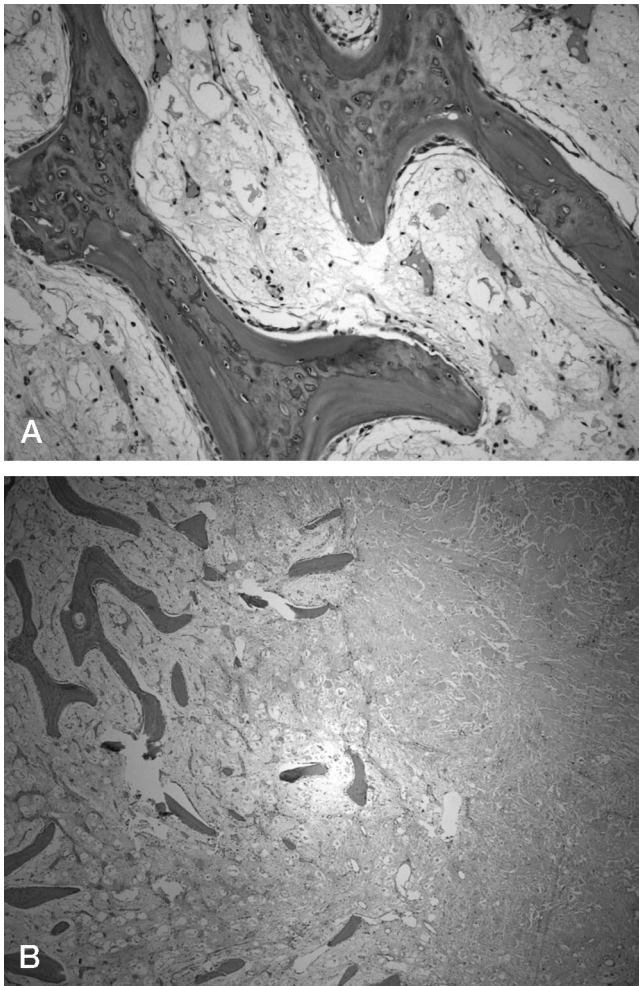


Figure 6. Hematoxylin and eosin slides of the center of the medullary canal of a BMP treated specimen. **A**, Trabeculae of immature new bone found within the medullary canal of the BMP treated group were denser at the host-graft interface (original magnification  $\times 200$ ). **B**, They thinned toward the edge of the central void (original magnification  $\times 40$ ).

tween groups was found during failure testing. Although the amount of torque at ultimate failure was more consistent in the BMP group than the autograft group (BMP average  $25.3 \pm 1.0$  Nm, range 24.7–27.3 vs. autograft average  $21.5 \pm 7.0$ , range 11.3–30.3), there was no statistical difference between groups ( $P = 0.22$ ).

### Complications

There were complications presumably related to the treatment in a total of 3 animals. As previously mentioned, one animal in the autograft group had a pseudarthrosis but survived 4 months with no noticeable functional deficit; the pseudarthrosis was detected only after harvest. The 2 additional animals were treated with BMP augmented grafts but were not included in the present analysis because they did not survive the full 4 months. Both animals went lame, and were euthanized 12 weeks after surgery and replaced. CT analysis of the lumbar spines revealed a clear pseudarthrosis in one animal but no bony abnormalities in the other animal. In this second animal, there was no canal or foraminal narrowing (by CT and gross examination) to explain the animal's refusal to bear weight. No cultures were taken in any of these animals to evaluate the presence of infection; however, there were no obvious abnormalities noted at necropsy.

### Discussion

Descriptions of the successful use of structural grafts in spinal reconstruction following corpectomy have all cited the importance of autograft augmentation of allografts; however, shortcomings of autograft procurement include volume restrictions and donor site morbidity. Augmentation of large allografts with rhBMP may provide equal fusion rates and fusion quality, while avoiding these downfalls. Although there are many studies to support the use of BMP in both spinal interbody fusions and

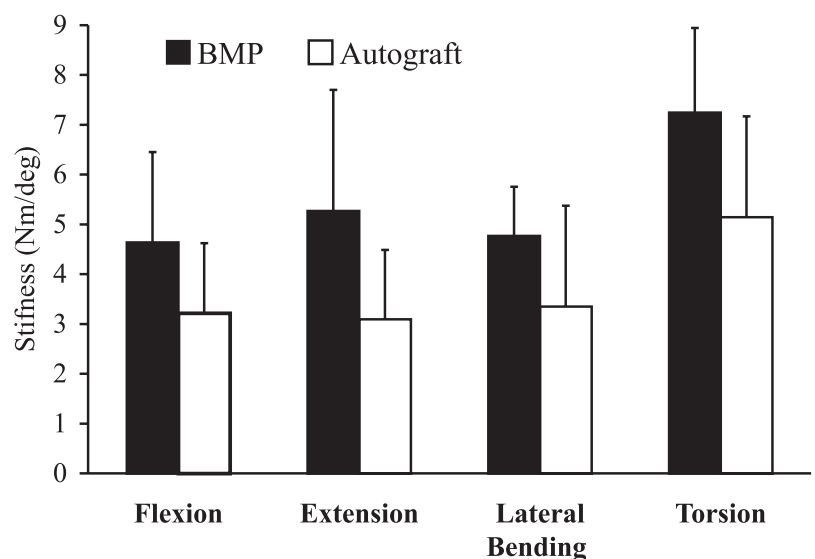


Figure 7. Biomechanical strength (mean  $\pm$  standard deviation) of allografts augmented with BMP show strength equivalent to allografts augmented with autograft in flexion ( $4.6 \pm 1.8$  vs.  $3.2 \pm 1.4$  Nm/deg,  $P = 0.16$ ), extension ( $5.3 \pm 2.4$  vs.  $3.1 \pm 1.4$  Nm/deg,  $P = 0.09$ ), lateral bending ( $4.8 \pm 1.0$  vs.  $3.4 \pm 2.0$  Nm/deg,  $P = 0.16$ ), and torsion ( $7.2 \pm 1.7$  vs.  $5.1 \pm 2.0$  Nm/deg,  $P = 0.14$ ).

the healing of large segment long bone defects, to the authors' knowledge, there are no reported studies of BMP augmentation of spinal reconstruction following corpectomy. The goal of this study was to determine if rhBMP-2 augmented strut grafts had similar fusion strength and graft incorporation to allografts augmented with autograft.

Radiographic, histologic, and biomechanical analyses revealed no significant differences between the BMP and autograft treatment groups. The radiographic analysis by CT revealed solid fusion in 7 of the 8 control animals (autograft) and 8 of the 8 experimental animals (BMP) that survived the full 4 months. Histologic analysis further elucidated the patterns of growth; samples from both groups had comparable levels of cortical strut incorporation. In the control histology specimen, a uniform distribution of remodeling trabeculae was found throughout the medullary canal; in contrast, the BMP specimens had regular trabecular bone formation at both host-bone interfaces, with a variably sized void at the center of the graft. A similar pattern of medullary bone was quantified by the CT analysis of the entire set of specimens. Biomechanical testing revealed similar stiffness in both nondestructive and failure testing modes. From this series, the authors surmise that in this bovine model, structural allograft augmented with BMP was equivalent to allograft augmented with autograft for the treatment of corpectomy defects.

### **Corpectomy and Strut Graft Fusion**

There have been several descriptions of radiographic outcomes in patients following corpectomy and strut graft fusion. Bridwell *et al*<sup>5</sup> studied 24 patients undergoing surgical correction for anterior column defects with anterior fresh frozen allograft supplemented by autogenous bone graft and stabilized with posterior instrumentation. The investigators found full radiographic fusion and trabecular remodeling in 18 of 24 patients, partial graft incorporation in 4 other patients, and emphasized the importance of rigid posterior instrumentation and autograft augmentation of the allograft. Similarly, Meding and Stambough<sup>24</sup> reviewed the outcomes of 51 patients after corpectomy and strut graft reconstruction. The fusion rates were determined clinically by the absence of pain and radiographically by evidence of healing, including "graft incorporation, callus formation, and bridging trabeculae at the strut graft-vertebral body interface." They found more than 90% fusion rates in both allograft and autograft treated groups but noted that fibular allografts appeared to take longer to incorporate than fibular autografts of a similar size.

Both of these studies involved a longer incorporation time than the survival period of the current study; however, this study showed fusion rates at this early time that were not only similar between treatment groups but also similar to those reported in these clinical studies. One of the 8 animals in the autograft group had pseudarthrosis, resulting in an 88% successful fusion rate in this group.

Nine of the 10 animals with BMP treatment had successful graft fusion (90%), including the 2 BMP animals that only survived 12 weeks (one with an obvious pseudarthrosis and one appearing united). Given the small numbers of specimens and that there was one clear pseudarthrosis in each group, the authors conclude that there was no significant difference in radiographic fusion rates between groups.

These rates are also similar to the fusion rates reported in other studies of BMP in spinal fusion. There are numerous experimental studies of BMP augmentation of posterolateral or anterior interbody fusion, with reported fusion rates ranging from 71% to 100%.<sup>13,16–18,27–29,33–37</sup> In one of the earliest studies of BMP treated spinal fusion, Lovell *et al*<sup>28</sup> used a dog posterior spinal fusion model and found a superior rate of fusion in vertebral levels with BMP than those with autograft alone. Cook *et al*<sup>27</sup> used a dog lumbar facet fusion model, and reported that both BMP and autograft augmented fusions had 100% fusion; however, the fusions assisted by BMP occurred in half the time of the autografted fusions (12 *vs.* 26 weeks). Both Boden<sup>13</sup> and Schimandle<sup>17</sup> *et al* evaluated posterolateral spinal fusions in a rabbit model and found a 100% fusion rate in BMP augmented fusions in contrast to significantly lower fusion rates in autograft controls. Studies of BMP augmentation in anterior interbody fusion in larger animal models have had higher fusion rates in BMP specimens (95% to 100%), using BMP concentrations similar to that used in the current study, than in autograft specimens.<sup>18,30,38</sup>

Several clinical studies have described the successful use of BMP in spinal fusion. All autograft control matched studies have shown equivalent or superior BMP fusion that ranges from 95% to 100%.<sup>14,39–45</sup> The largest clinical trial published to date, a multicenter collaboration by Burkus *et al*,<sup>41</sup> analyzed 279 patients who underwent lumbar interbody fusion using an LT-CAGE® tapered fusion device (Medtronic Sofamor Danek), with either autograft or BMP soaked collagen sponges inside the device. At 2-year follow-up, the BMP group had a 95% fusion rate, slightly higher than the control fusion rate of 89% ( $P = 0.022$ ). None of these studies of BMP in spinal fusion (experimental or clinical) use a corpectomy model to analyze the efficacy of BMP in the incorporation of a large structural allograft. Some extrapolation from studies of BMP treatment of large segment long bone defects can be made; however, to the authors' knowledge, this study represents the first examination of BMP augmentation of large structural allograft incorporation in the spine.

### **Graft Incorporation**

Despite the finding of equivalent fusion rates between the experimental and control group, CT and histologic analyses revealed different patterns of overall strut graft incorporation. There was no significant difference between groups regarding resorption and creeping substitution found in the cortex of the graft. Image analysis of the



residual cortex showed similar rates of resorption between groups, with more resorption at the host-graft interface. At this interface in all histology specimens, there was organized and dense trabecular bone. This pattern is similar to large graft incorporation patterns reported in both the experimental and clinical studies in which incorporation begins at the host-graft interface and extends inward along the osteoconductive lattice of the allograft.<sup>5,24,25,46–48</sup> However, the graft may never fully repair; necrotic bone may persist, imbedded in the remodeled live bone of a successful fusion. The evaluation of retrieved large allografts by Enneking and Campanacci<sup>47</sup> showed that these allografts anneal and incorporate at the ends, and along the periphery of the graft with only approximately 15% to 20% of the total allograft volume actually remodeling. Longer survival times may show more complete graft cortex remodeling in the treatment groups; however, successful fusion does not depend on complete remodeling of the strut.

Other differences in graft cortex remodeling may have been seen if the survival time had been longer. In a study of rhBMP-2 treatment of dog segmental femoral defects by Zabka *et al*,<sup>22</sup> the investigators compared allograft incorporation in specimens augmented with autograft, BMP infused collagen sponges, and collagen sponges alone. They reported higher porosity and pore diameters of the remodeling allograft cortex in the control groups as compared to the BMP group, and concluded that there was less resorption in the BMP specimens. The investigators also found increased bone formation in the BMP specimens and deduced there was more “balanced bone resorption/formation in allograft repair induced by rhBMP-2.” This more balanced creeping substitution could result in earlier and stronger fusion of the overall construct.

The major difference in graft incorporation between groups was found in the medullary canal of the strut graft. At the index procedure, the control grafts were filled with local cancellous autograft. At harvest, neovascularization and islands of appositional bone throughout the length of the canal indicated that incorporation of the autograft was in progress. The individual trabeculae contained both necrotic bone with empty lacunae, cement lines, and adjacent immature woven bone, indications of active bone remodeling as originally described by Urist.<sup>49</sup> These findings are similar to the histologic findings of Togawa *et al*.<sup>26</sup> On histologic analysis of needle biopsies taken from within successfully fused human intervertebral body cages, the investigators found variable degrees of necrotic bone as many as 6 years after the initial surgery. These bone islands, remodeling remnants of the autograft, were not as organized or interconnected as the bone at the host-graft interface or even normal cancellous bone. Once again, a longer survival period would be needed to ascertain the complete organization of the new bone within the graft canal.

In the BMP group, CT revealed dense and organized cancellous bone at the distal and proximal ends of the graft. However, in the middle section of the graft, there was a variably sized void of mineralization. This pattern suggests graft incorporation occurred mainly on the graft ends. No definitive conclusions can be made about the fate of the central void. The histologic findings of vessel canals within the central void, as well as thinner spicules at the interface between the void and the newly formed bone, suggest that this is an advancing edge of osteogenesis. However, a higher number of specimens and studies of a longer survival period would be necessary to ascertain how much of the medullary canal could be filled by BMP initiated bone formation.

In a study of the treatment of large segmental defects in dog ulnae, Salkeld *et al*,<sup>21</sup> witnessed the successful healing of 2.5-cm gaps in 5 out of 5 animals treated solely with a collagen/BMP device. In the present study, the defect was approximately 3.6 cm; however, the osteoconductive lattice of the allograft may augment fusion and allow the BMP instigated cascade of bone formation to bridge this larger gap and continue until the void is filled. On the contrary, the presence of a stable fusion may shield the bone formation process from stress, causing osteogenesis to slow and halt before filling the void. Bouxsein *et al*<sup>50</sup> found that only approximately 8% of BMP remains in collagen sponges at 2 weeks after implantation and theorized that the initial presence of BMP begins a cascade of bone formation that continues after elution of the growth factor. However, after osteoinduction, bone formation is governed by the fundamental rules of healing, and depends on the demands and resources of the host environment.

### Study Limitations

There were several limitations to the model used in this study. The primary limitation was the use of an immature bovine model. Familiarity with the calf spinal model and appropriate size for instrumentation had its advantages. However, as an immature animal, the capacity for healing and fusion was potentially higher, and not necessarily comparable to the typical patient undergoing corpectomy and reconstruction. Although the fusion rates were similar to interbody fusion in other animal models for both autograft and BMP augmentation, the pseudarthrosis rate in this large graft incorporation may have been underrepresented and obscured by the substantial healing potential of this immature model. Although large animals such as goats and sheep have been used, bovines have not been used as an animal model for BMP spinal studies. However, it is likely that these animals of similar phylogenetic order have similar BMP dose responses.

Another limitation of this study was the small sample size. Larger animals, although necessary for the size of their spine in the experimental procedure, are resource intensive animals, keeping the study population small and making statistical power difficult to achieve. In the

study design, only 2 animals per group were designated for histologic evaluation, leaving the rest for definitive analysis of the fusion strength with biomechanical testing. Despite the fact that there was only one autograft specimen and 2 BMP specimens evaluated, the histologic findings were well supported by existing literature on BMP osteoinduction and graft incorporation. In addition, distribution of bone in these specimens was exactly echoed in the CT studies performed on these specimens before histologic preparation. Consequently, the histologic findings of bone growth, organization, and resorption can be extrapolated to the other specimens.

Finally, the present study is only the first step in the investigation of BMP augmentation of large structural allografts in the spine. The concentration of BMP used in this study was taken from an interbody fusion study conducted in sheep.<sup>30</sup> However, it is well known that fusion with BMP requires lower concentrations in nonprimate models.<sup>36</sup> Additional testing in a nonhuman primate model or extrapolation from concentrations used in human interbody fusions would be necessary to provide a clinically effective concentration guideline. Furthermore, the use of a mature nonhuman primate model would better approximate the adult human condition, in which this type of fusion is typically more difficult to achieve.

## ■ Conclusion

This study provided a description of the process by which early healing and incorporation of large spinal cortical strut allografts occurs. Additionally, this study of a lumbar corpectomy model showed that cortical allograft struts supplemented with rhBMP-2 fused with strength equivalent to allografts filled with cancellous autograft. These results, similar to those previously seen with shorter intradiscal grafts/spacers, suggest the efficacy of BMP in large segment reconstructions after vertebral corpectomy. The longer segment of allograft required for corpectomy reconstruction provided a more demanding environment for induction of new bone formation by the BMP compared to the intradiscal grafts studied in the past. During the 4-month survival period of this study, new bone formation within the allograft was incomplete in the central segments of the allograft strut in both treatment groups. However, there was evidence of organizing bony growth in the autograft group and an advancing front of the vascular ingrowth with trailing new bone formation in the BMP group.

When BMP were first discovered through the work of Urist,<sup>49</sup> the hope began that eventually patients would be spared the morbidity of autograft harvesting, and surgeons would have an unlimited resource of osteoinductive elements. Although future studies will be required to address the issues of long-term graft remodeling, human dosage, and clinical outcomes, this study illustrates the feasibility of BMP augmentation for the healing of large segment allografts in the spine.

## ■ Key Points

- There was no significant difference in the rate of fusion for structural allografts augmented with rhBMP-2 *versus* autograft.
- Fusion in BMP treated specimens was as strong as fusion created by the gold standard autograft filled strut graft.
- Despite differences in new bone distribution, there was no significant difference in allograft incorporation between experimental groups.
- The rhBMP-2 may be a useful adjunct in the placement of long structural allografts in the spine.

## References

1. Esses SI, Botsford DJ, Kostuik JP. Evaluation of surgical treatment for burst fractures. *Spine* 1990;15:667-73.
2. Harrington KD. Anterior decompression and stabilization of the spine as a treatment for vertebral collapse and spinal cord compression from metastatic malignancy. *Clin Orthop Relat Res* 1988;233:177-97.
3. Hodgson AR, Stock FE. Anterior spinal fusion. A preliminary communication on the radical treatment of Pott's disease and Pott's paraplegia. 1956. *Clin Orthop Relat Res* 1994;300:16-23.
4. Younger EM, Chapman MW. Morbidity at bone graft donor sites. *J Orthop Trauma* 1989;3:192-5.
5. Bridwell KH, Lenke LG, McEnery KW, et al. Anterior fresh frozen structural allografts in the thoracic and lumbar spine. Do they work if combined with posterior fusion and instrumentation in adult patients with kyphosis or anterior column defects? *Spine* 1995;20:1410-8.
6. Vail TP, Urbaniak JR. Donor-site morbidity with use of vascularized autogenous fibular grafts. *J Bone Joint Surg Am* 1996;78:204-11.
7. Robertson PA, Wray AC. Natural history of posterior iliac crest bone graft donation for spinal surgery: A prospective analysis of morbidity. *Spine* 2001;26:1473-6.
8. Singh K, DeWald CJ, Hammerberg KW, et al. Long structural allografts in the treatment of anterior spinal column defects. *Clin Orthop Relat Res* 2002;394:121-9.
9. An HS, Simpson JM, Glover JM, et al. Comparison between allograft plus demineralized bone matrix versus autograft in anterior cervical fusion. A prospective multicenter study. *Spine* 1995;20:2211-6.
10. Wimmer C, Krismer M, Gluch H, et al. Autogenic versus allogenic bone grafts in anterior lumbar interbody fusion. *Clin Orthop Relat Res* 1999;360:122-6.
11. Cohen DB, Chotivichit A, Fujita T, et al. Pseudarthrosis repair. Autogenous iliac crest versus femoral ring allograft. *Clin Orthop Relat Res* 2000;371:46-55.
12. Gibson S, McLeod I, Wardlaw D, et al. Allograft versus autograft in instrumented posterolateral lumbar spinal fusion: A randomized control trial. *Spine* 2002;27:1599-603.
13. Boden SD, Schimandle JH, Hutton WC. Lumbar intertransverse-process spinal arthrodesis with use of a bovine bone-derived osteoinductive protein. A preliminary report. *J Bone Joint Surg Am* 1995;77:1404-17.
14. Boden SD, Zdeblick TA, Sandhu HS, et al. The use of rhBMP-2 in interbody fusion cages. Definitive evidence of osteoinduction in humans: A preliminary report. *Spine* 2000;25:376-81.
15. Hecht BP, Fischgrund JS, Herkowitz HN, et al. The use of recombinant human bone morphogenetic protein 2 (rhBMP-2) to promote spinal fusion in a nonhuman primate anterior interbody fusion model. *Spine* 1999;24:629-36.
16. Muschler GF, Hyodo A, Manning T, et al. Evaluation of human bone morphogenetic protein 2 in a canine spinal fusion model. *Clin Orthop Relat Res* 1994;308:229-40.
17. Schimandle JH, Boden SD, Hutton WC. Experimental spinal fusion with recombinant human bone morphogenetic protein-2. *Spine* 1995;20:1326-37.
18. Zdeblick TA, Ghanayem AJ, Rapoff AJ, et al. Cervical interbody fusion cages. An animal model with and without bone morphogenetic protein. *Spine* 1998;23:758-65.
19. Friedlaender GE, Perry CR, Cole JD, et al. Osteogenic protein-1 (bone morphogenetic protein-7) in the treatment of tibial nonunions. *J Bone Joint Surg Am* 2001;83-A:S151-8.
20. Pluhar GE, Manley PA, Heiner JPJR, et al. The effect of recombinant

- human bone morphogenetic protein-2 on femoral reconstruction with an intercalary allograft in a dog model. *J Orthop Res* 2001;19:308–17.
21. Salkeld SL, Patron LP, Barrack RL, et al. The effect of osteogenic protein-1 on the healing of segmental bone defects treated with autograft or allograft bone. *J Bone Joint Surg Am* 2001;83-A:803–16.
  22. Zabka AG, Pluhar GE, Edwards RB 3rd, et al. Histomorphometric description of allograft bone remodeling and union in a canine segmental femoral defect model: A comparison of rhBMP-2, cancellous bone graft, and absorbable collagen sponge. *J Orthop Res* 2001;19:318–27.
  23. Molinari RW, Bridwell KH, Klepps SJ, et al. Minimum 5-year follow-up of anterior column structural allografts in the thoracic and lumbar spine. *Spine* 1999;24:967–72.
  24. Meding JB, Stambough JL. Critical analysis of strut grafts in anterior spinal fusions. *J Spinal Disord* 1993;6:166–74.
  25. Enneking WF, Mindell ER. Observations on massive retrieved human allografts. *J Bone Joint Surg Am* 1991;73:1123–42.
  26. Togawa D, Bauer TW, Brantigan JW, et al. Bone graft incorporation in radiographically successful human intervertebral body fusion cages. *Spine* 2001;26:2744–50.
  27. Cook SD, Dalton JE, Tan EH, et al. In vivo evaluation of recombinant human osteogenic protein (rhOP-1) implants as a bone graft substitute for spinal fusions. *Spine* 1994;19:1655–63.
  28. Lovell TP, Dawson EG, Nilsson OS, et al. Augmentation of spinal fusion with bone morphogenetic protein in dogs. *Clin Orthop Relat Res* 1989;243:266–74.
  29. Sandhu HS, Kanim LE, Kabo JM, et al. Evaluation of rhBMP-2 with an OPLA carrier in a canine posterolateral (transverse process) spinal fusion model. *Spine* 1995;20:2669–82.
  30. Sandhu HS, Toth JM, Diwan AD, et al. Histologic evaluation of the efficacy of rhBMP-2 compared with autograft bone in sheep spinal anterior interbody fusion. *Spine* 2002;27:567–75.
  31. Sanderson C, Kitabayashi LR. Parallel experience of two different laboratories with the initiator perkadox 16 for polymerization of methylmethacrylates. *J Histotechnol* 1994;17:343–8.
  32. Sanderson C, McGee M, Bloebaum RD. Polypropylene containers for safe and predictable embedding of specimens in PMMA. *J Histotechnol* 1990;13:131–3.
  33. Akamaru T, Suh D, Boden SD, et al. Simple carrier matrix modifications can enhance delivery of recombinant human bone morphogenetic protein-2 for posterolateral spine fusion. *Spine* 2003;28:429–34.
  34. Boden SD, Martin GJ Jr, Morone MA, et al. Posterolateral lumbar intertransverse process spine arthrodesis with recombinant human bone morphogenetic protein 2/hydroxyapatite-tricalcium phosphate after laminectomy in the nonhuman primate. *Spine* 1999;24:1179–85.
  35. Cunningham BW, Kanayama M, Parker LM, et al. Osteogenic protein versus autologous interbody arthrodesis in the sheep thoracic spine. A comparative endoscopic study using the Bagby and Kuslich interbody fusion device. *Spine* 1999;24:509–18.
  36. Martin GJ Jr, Boden SD, Marone MA, et al. Posterolateral intertransverse process spinal arthrodesis with rhBMP-2 in a nonhuman primate: Important lessons learned regarding dose, carrier, and safety. *J Spinal Disord* 1999;12:179–86.
  37. Sandhu HS, Kanim LE, Kabo JM, et al. Effective doses of recombinant human bone morphogenetic protein-2 in experimental spinal fusion. *Spine* 1996;21:2115–22.
  38. Sucato DJ, Hedequist D, Zhang H, et al. Recombinant human bone morphogenetic protein-2 enhances anterior spinal fusion in a thoracoscopically instrumented animal model. *J Bone Joint Surg Am* 2004;86:752–62.
  39. Boden SD, Kang J, Sandhu H, et al. Use of recombinant human bone morphogenetic protein-2 to achieve posterolateral lumbar spine fusion in humans: A prospective, randomized clinical pilot trial: 2002 Volvo Award in clinical studies. *Spine* 2002;27:2662–73.
  40. Burkus JK, Dorchak JD, Sanders DL. Radiographic assessment of interbody fusion using recombinant human bone morphogenetic protein type 2. *Spine* 2003;28:372–7.
  41. Burkus JK, Gornet MF, Dickman CA, et al. Anterior lumbar interbody fusion using rhBMP-2 with tapered interbody cages. *J Spinal Disord Tech* 2002;15:337–49.
  42. Burkus JK, Transfeldt EE, Kitchel SH, et al. Clinical and radiographic outcomes of anterior lumbar interbody fusion using recombinant human bone morphogenetic protein-2. *Spine* 2002;27:2396–408.
  43. Baskin DS, Ryan P, Sonntag V, et al. A prospective, randomized, controlled cervical fusion study using recombinant human bone morphogenetic protein-2 with the CORNERSTONE-SR allograft ring and the ATLANTIS anterior cervical plate. *Spine* 2003;28:1219–25.
  44. Johnsson R, Stromqvist B, Aspenberg P. Randomized radiostereometric study comparing osteogenic protein-1 (BMP-7) and autograft bone in human noninstrumented posterolateral lumbar fusion: 2002 Volvo Award in clinical studies. *Spine* 2002;27:2654–61.
  45. Vaccaro AR, Patel T, Fischgrund J, et al. A pilot safety and efficacy study of OP-1 putty (rhBMP-7) as an adjunct to iliac crest autograft in posterolateral lumbar fusions. *Eur Spine J* 2003;12:495–500.
  46. Bos GD, Goldberg VM, Powell AE, et al. The effect of histocompatibility matching on canine frozen bone allografts. *J Bone Joint Surg Am* 1983;65:89–96.
  47. Enneking WF, Campanacci DA. Retrieved human allografts: A clinicopathological study. *J Bone Joint Surg Am* 2001;83-A:971–86.
  48. Dell PC, Burchardt H, Glowczewskie FP Jr. A roentgenographic, biomechanical, and histological evaluation of vascularized and non-vascularized segmental fibular canine autografts. *J Bone Joint Surg Am* 1985;67:105–12.
  49. Urist MR. Bone: Formation by autoinduction. 1965. *Clin Orthop Relat Res* 2002;395:4–10.
  50. Bouxsein ML, Turek TJ, Blake CA, et al. Recombinant human bone morphogenetic protein-2 accelerates healing in a rabbit ulnar osteotomy model. *J Bone Joint Surg Am* 2001;83-A:1219–30.



# Biomechanical Comparison of Lumbosacral Fixation Using Luque-Galveston and Colorado II Sacropelvic Fixation: Advantage of Using Locked Proximal Fixation

Sean Early, MD,\* Andrew Mahar, MS,†‡ Richard Oka, BA,† and Peter Newton, MD†‡

**Study Design.** Biomechanical evaluation of sacropelvic fixation strategies as they apply to neuromuscular scoliosis.

**Objectives.** The primary objective was to compare the rigidity of 2 methods of sacropelvic fixation (Galveston vs. Colorado II). The secondary objective was to evaluate the effect on construct rigidity by adding a pair of L1 pedicle screws to a Luque wire construct.

**Summary of Background Data.** The Galveston modification to the Luque rodding system has become standard for treating childhood and adolescent neuromuscular scoliosis. The Galveston method provides reasonable lumbo-pelvic fixation with a relatively simple method of insertion. Clinical reviews of sagittal plane stability in neuromuscular patients with Galveston fixations performed at our institution have led to concerns regarding the technique's ability to maintain proper lumbar lordosis. This concern has generated interest in evaluating biomechanical stability of more rigid fixation methods in these long spino-pelvic constructs. As such, the following biomechanical study evaluated lumbosacral stability of 2 sacropelvic fixation methods: the standard Luque-Galveston method and the Colorado II sacropelvic fixation method using the Chopin plate-screw block. As a secondary interest, evaluations of the rigidity of the proximal construct when using pedicle screw fixation were completed. It was hypothesized that one additional point of rigid fixation at the thoracolumbar junction may make substantial improvement in rigidity to the otherwise Luque construct.

**Methods.** Lumbo-pelvic segments of human cadaveric specimens were instrumented with L1 pedicle screws, sublaminar wires between L2 and L5, and sacropelvic fixation with either Galveston rods or Colorado II sacropelvic plates using S1 screws, S2 alar screws, and iliac screws. Tests were conducted for physiologic flexion-extension and torsional loading. Construct stiffness be-

tween L1–S1 was determined for each specimen. Motion measurement data were collected between L1–L5 and L5–S1 using a noncontact marker system. Statistical analysis included a 2-way analysis of variance (dependent variables: construct/locked screw) with the Tukey *post hoc* correction for multiple comparisons ( $P < 0.05$ ).

**Results.** The flexion and extension bending stiffness of the construct was similar between the Galveston and Colorado II constructs. Both constructs were stiffer when the L1 screws were locked rigidly to the rod. Torsional stiffness followed similar trends with no significant difference between the systems, although more rigid in more cases when the L1 screws were locked to the rod. Regarding limiting L5–S1 motion during flexion and extension loading, the Colorado II construct did so to a higher degree. There was no difference in torsional motion between the 2 constructs. Locking the L1 pedicle screws reduced torsional motion but had no effect on flexion-extension motion at L5–S1.

**Conclusion.** The 2 methods of sacropelvic fixation provided similar construct stiffness, although the Colorado II method had less L5–S1 motion on flexion-extension testing, and the Galveston construct tended (although not statistically) to be stiffer in torsional loading. The addition of a pair of L1 pedicle screws increased the construct stiffness for both constructs by approximately 50%.

**Key words:** neuromuscular scoliosis, lumbosacral fixation, fusion, vertebral motion, biomechanics. **Spine** 2005;30:1396–1401

Great advances have been made in the surgical treatment of neuromuscular scoliosis. However, challenges remain, particularly with fixation at the lumbosacral junction. As a result of the long length of fusion and the substantial lever arm, large moments are placed at the lumbosacral junction. Secure fixation at the sacropelvic foundation of the construct is required to resist these loads.

During the initial development of spinal instrumentation for neuromuscular patients, Eduardo Luque popularized the use of multiple sublaminar wires for segmental fixation at each level.<sup>1,2</sup> This provides a relatively simple method for attaining gradual correction of the deformity by allowing the wires to be sequentially tightened. However, pelvic fixation was less than ideal with the distal ends of the rod bent into an L-shape for transiliac fixation. Modifications to this form of distal fixation resulted from the work of Allen and Ferguson<sup>3–5</sup> with the creation of the Galveston method of distal rod insertion. This method involves a series of bends to the distal rod, which allow insertion of the end of the rod through a column of bone in the posterior ilium, just

From the \* Division of Orthopaedic Surgery, Children's Hospital-Los Angeles, Los Angeles, CA, †Department of Orthopedics, Children's Hospital-San Diego, San Diego, CA, and ‡Department of Orthopedics, University of California, San Diego, San Diego, CA.

Acknowledgment date: September 23, 2004. First revision date: June 2, 2004. Acceptance date: July 23, 2004.

Supported by a Medtronic Sofamor Danek research grant.

The device(s)/drug(s) is/are FDA-approved or approved by corresponding national agency for this indication.

Corporate/Industry funds were received in support of this work. Although one or more of the author(s) has/have received or will receive benefits for personal or professional use from a commercial party related directly or indirectly to the subject of this manuscript, benefits will be directed solely to a research fund, foundation, educational institution, or other nonprofit organization which the author(s) has/have been associated.

Address correspondence and reprint requests to Peter Newton, MD, 3020 Children's Way, Suite 410, San Diego, CA 92123-4293; E-mail: pnewton@chsd.org

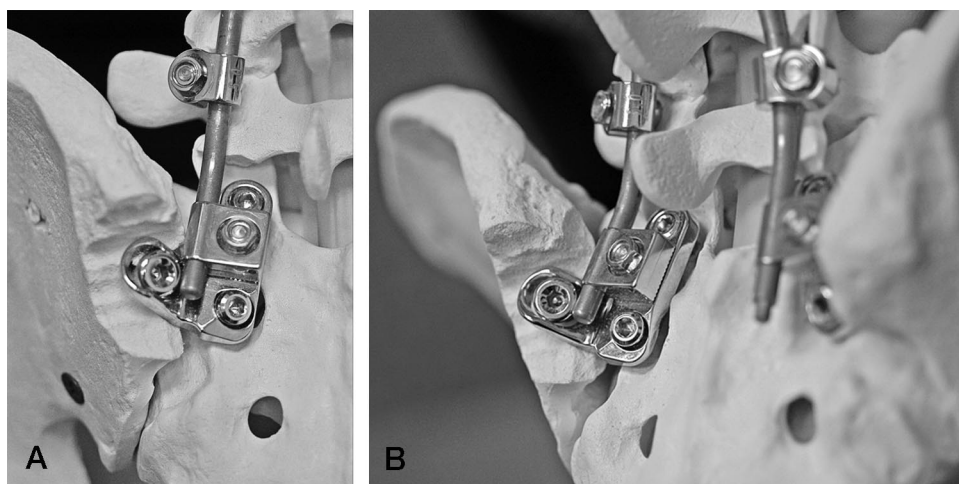


Figure 1. **A** and **B**, Chopin block instrumentation.

superior to the greater sciatic notch. This form of fixation was refined with the development of the “unit rod.” This precontoured rod provides the required bends for iliac fixation, as well as precontoured lumbar lordosis and thoracic kyphosis in a single U-shaped rod.<sup>6,7</sup> This rod has been used widely for neuromuscular scoliosis reconstruction and remains one of the most common methods of instrumentation for such curves.<sup>7–14</sup>

The segmental fixation involves Luque sublaminar wires, which may be combined or replaced by hooks and/or screws, such as those used in idiopathic scoliosis. The Luque method does not provide completely rigid axial fixation because the spine may telescope up and down the rods. Torsional fixation is also less than rigid, although it can be enhanced by using rod cross-connecting devices.<sup>12</sup> This lack of rigidity may limit the ability of the Luque-Galveston method to maintain deformity correction, especially at the lumbosacral junction.

Although much focus has been placed on the coronal plane deformity (scoliosis and pelvic obliquity) in these patients, sagittal plane malalignment (kyphosis, lordosis) may also result in substantial difficulties in sitting.<sup>13</sup> Although the Galveston construct in the sacropelvis appears to be reasonably secure in the coronal plane,<sup>14</sup> it is less so in the sagittal plane. With flexion at the lumbosacral disc, the rods may pull away from the pelvis and become loose and/or prominent distally.

Because of these perceived limitations to the Galveston method of sacropelvic fixation, an investigation was undertaken to assess another method of sacropelvic fixation using a plate-screw combination.<sup>15</sup> The “Chopin block” with 2 sacral screws and one iliac screw (Figure 1) was compared to the Galveston method regarding acute construct stability. As a secondary interest, the influence of a single level of rigid proximal fixation (a pair of L1 pedicle screws) on the stability of both constructs was also evaluated as a possible way for increasing lumbosacral construct rigidity.

## ■ Methods

Eight cadaveric human lumbar spine-sacropelvic specimens were obtained and stored frozen. Before testing, the muscular

tissues were removed, with the ligaments maintained. The end of the sacrum was potted into 2-part epoxy resin and bolted into a fixture. The proximal spine was also potted in epoxy resin and attached with screws, either to a cantilever apparatus for applying bending moments or within a potting fixture for applying torque.

The intact specimen was then mounted onto the MTS 858 servo-hydraulic testing machine (MTS Systems Corp., Eden Prairie, MN). Torsional testing in axial rotation was performed at 1° per second, with an upper torque limit of  $\pm 2$  Nm. Angular displacement and torque data were collected at 20 Hz for 20 cycles. Stiffness data were calculated for the last 4 cycles and averaged for each specimen. Subsequently, the upper mount of the spine was changed to a cantilever bending apparatus, and bending moments of 5 Nm in flexion and extension at approximately 2° per second were applied. Flexion and extension tests were performed for 20 cycles. Based on the cantilever bending apparatus dimensions, the axial load data were converted to torque. Stiffness data were calculated for the last 4 cycles and averaged for each specimen. Motion data were recorded using reflective markers and a Qualisys motion measurement system (Qualisys Medical AB, Gothenburg, Sweden) sampling at 20 Hz over the last 4 cycles. The marker clusters were placed on L1, L5, and S1. The order of torsional and flexion-extension testing was randomized for all specimens.

After mechanical testing of the intact spine, the specimens were instrumented from L1 to the pelvis. The intraspinal ligaments were removed to allow passage of 18-gauge sublaminar wires into the vertebral canal at L2, L3, L4, and L5. At the L1 level, polyaxial 6.5-mm diameter pedicle screws were also placed. A Luque-Galveston construct was created using a 5.5-mm diameter unit rod, and cross links were added at L2 and L5 for the Luque construct only (Figure 2). Identical testing, as described previously for the intact specimen, was repeated for the Galveston construct. The mechanical testing sequence was performed twice for each Luque-Galveston construct, once with L1 pedicle screws tightly locked to the rods and once with the screws engaging the rods but not locked. The unlocked condition simulated wire fixation at that level, allowing some degree of axial movement along the length of the rod as occurs with Luque instrumentation. This is in contrast to the locked condition in which this sliding and length change is restricted. Following that testing, the specimen was removed from the MTS machine and prepared for the Chopin block construct.

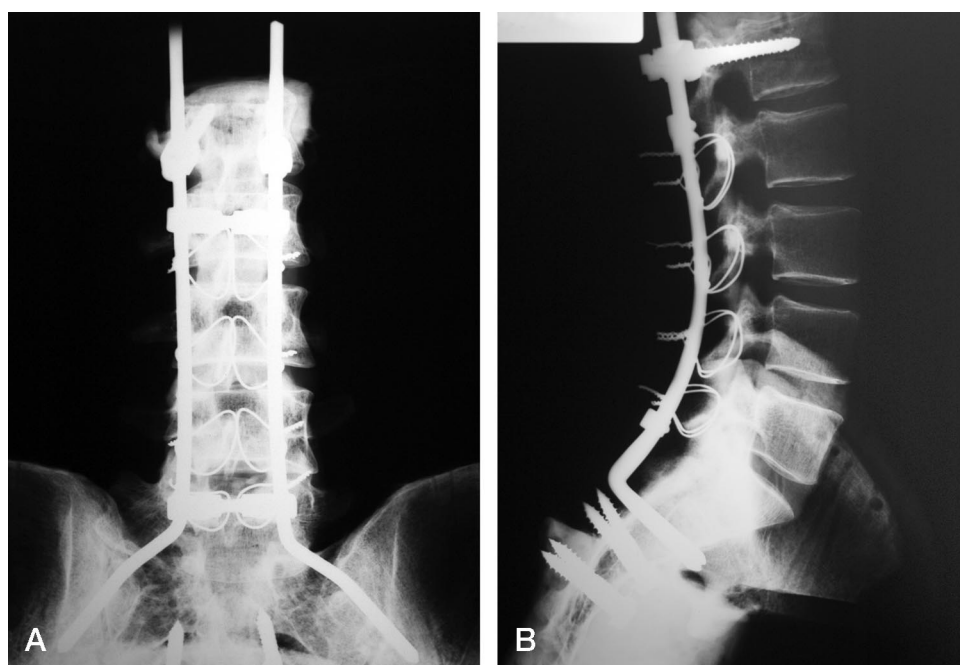


Figure 2. Anteroposterior (A) and lateral (B) radiographs of a specimen instrumented with the Luque-Galveston construct.

The Colorado II sacropelvic instrumentation (Medtronic Sofamor-Danek, Memphis, TN) was applied using S1 pedicle and S2 alar screws, as well as an iliac screw placed through a 3-hole Chopin block. A similarly contoured unit rod was cut off just above the lower bend, connected to the Colorado II sacral plates, and secured with new sublaminar wires (Figure 3). Motion data were acquired as before, both with the L1 screws locked as well as engaged but not tight (unlocked). Thus, each spinal pelvic specimen was instrumented with both methods and tested in the same manner. Although the order of load testing (torsion/flexion-extension) was randomized, the Galveston testing always preceded the Colorado II testing because the Galveston rod (inserted into the ilium) was smaller in diameter (5.5 mm) than the iliac screw placed in the same path

of the Colorado II construct (7.0 mm). Each specimen was radiographed following instrumentation to ensure proper placement of each type of fixation system.

Mechanical data for each test (angular rotation, torque) were collected, and stiffness was calculated from the linear portion of the torque-rotation curve (0.5–2 Nm for torsion, 0.5–5 Nm for flexion-extension). Differences in L5–S1 motion were also evaluated for both axial rotation and flexion-extension bending. Data for bending and torsional construct stiffness were analyzed using a 2-way analysis of variance (ANOVA) ( $P < 0.05$ ), with independent variables being the type of sacropelvic fixation and use of locked/unlocked L1 pedicle screws. Tukey *post hoc* testing was performed when significant differences were found to perform multiple comparisons as indicated.

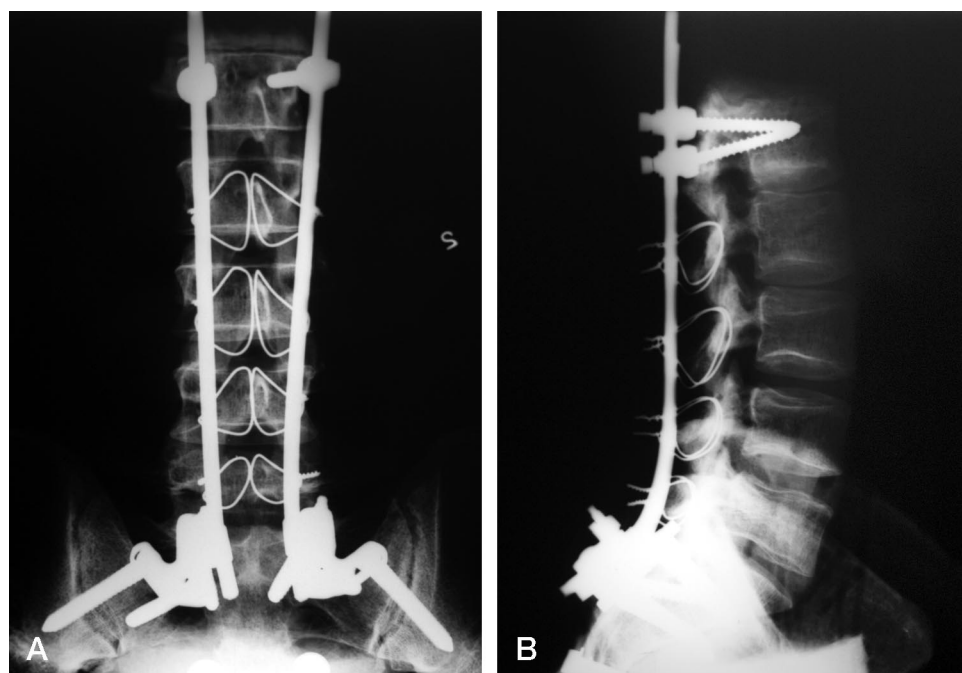


Figure 3. Anteroposterior (A) and lateral (B) radiographs of a Colorado II sacropelvic fixation.



**Table 1. 2-Way ANOVA Results for L1– S1 Flexion-Extension Stiffness (Nm/deg)**

	Mean $\pm$ SD		P Value
Colorado			
Locked	1.38 $\pm$ 0.19	Colorado vs. Galveston	0.97
Unlocked	0.93 $\pm$ 0.24		
Galveston			
Locked	1.45 $\pm$ 0.49	Locked vs. unlocked	0.03
Unlocked	0.89 $\pm$ 0.14		
Intact spine	0.34 $\pm$ 0.19	Interaction term	0.56

## ■ Results

### L1–S1 Stiffness

The overall construct stiffness differed very little between the 2 methods of sacropelvic fixation. In flexion-extending bending, the stiffness was  $0.9 \pm 0.1$  Nm/deg and  $0.9 \pm 0.2$  Nm/deg for the Galveston and Colorado II constructs, respectively, when the L1 screws were loose-unlocked. Each also resulted in similar flexion-extension stiffness when the L1 screws were locked tightly to the rods ( $1.5 \pm 0.5$  vs.  $1.4 \pm 0.1$ , respectively), although both constructs were stiffer with the L1 screws locked. The 2-way ANOVA confirmed this statistically with no difference between the Galveston and Colorado II constructs in flexion-extension stiffness ( $P > 0.9$ ); however, the addition of L1 screws did statistically increase the stiffness ( $P = 0.03$ ). The interaction term was 0.5 (Table 1).

Torsional stiffness of the instrumented construct (between L1 and S1) was slightly but not statistically higher for the Galveston construct compared to the Colorado II system ( $P = 0.3$ ). However, a statistical effect from locking the L1 screws was identified (*i.e.*, stiffer with the screws locked) ( $P < 0.005$ ). The torsional stiffness of the Galveston constructs was  $2.5 \pm 0.6$  Nm/deg and  $1.4 \pm 0.4$  Nm/deg for the locked and unlocked L1 screw condition, respectively. This result compares to  $2.0 \pm 0.4$  Nm/deg and  $1.2 \pm 0.4$  Nm/deg for the respective Colorado II constructs (Table 2).

### L5–S1 Motion

Regarding limiting L5–S1 motion during flexion-extension loading, the Colorado II construct limited motion to a higher degree ( $P < 0.001$ ). The total motion was  $3.2^\circ \pm 0.8^\circ$  and  $2.8^\circ \pm 0.7^\circ$  for the locked and unlocked

**Table 3. 2-Way ANOVA Results for L5– S1 Flexion-Extension Motion (degrees)**

	Mean $\pm$ SD		P Value
Colorado			
Locked	3.21 $\pm$ 0.84	Colorado vs. Galveston	0.001
Unlocked	2.78 $\pm$ 0.70		
Galveston			
Locked	4.22 $\pm$ 1.18	Locked vs. unlocked	0.78
Unlocked	4.46 $\pm$ 1.14		
Intact spine	8.73 $\pm$ 2.76	Interaction term	0.34

L1 screw condition, respectively. This result compares to  $4.2^\circ \pm 1.2^\circ$  and  $4.9^\circ \pm 1.1^\circ$  for the 2 respective Galveston constructs. Locking the L1 pedicle screws did not affect flexion-extension motion between L5 and S1 ( $P = 0.8$ ) (Table 3). Torsional motion at L5–S1 was not satisfactorily different between the systems, although there was a trend toward reduced torsional motion with the Galveston construct ( $1.2^\circ \pm 0.4^\circ$  locked and  $1.4^\circ \pm 0.7^\circ$  unlocked) compared to the Colorado II system ( $2.0^\circ \pm 0.8^\circ$  locked and  $2.4^\circ \pm 1.0^\circ$  unlocked). As opposed to the flexion-extension testing, the 2-way ANOVA suggested less torsional motion when the L1 pedicle screws were tight to the rods ( $P = 0.04$ ) (Table 4).

## ■ Discussion

Fixation across the lumbosacral junction remains a challenge, particularly for correction of neuromuscular spinal deformity.<sup>16–22</sup> Correction and stability in the axial, coronal, and sagittal planes is required. The Luque-Galveston method has provided a satisfactory solution, yet room for improvement clearly exists. Loosening of the pelvic fixation with backing out of the rods is problematic in a percentage of patients. The cases that result in the greatest challenge with the Galveston method are often associated with marked sagittal plane deformities (*i.e.*, both kyphosis and lordosis). In cases of hyperlordosis, it is difficult to achieve the appropriate angle of insertion of the iliac portion of a Galveston rod, especially if it is a unit rod. In cases of hyperkyphosis, the Galveston rod is angled anteriorly, with little ability to resist flexion and maintain an extended posture of the lumbar spine relative to the pelvis.

The plate and screw system of sacropelvic fixation

**Table 2. 2-Way ANOVA Results for L1– S1 Torsion Stiffness (Nm/deg)**

	Mean $\pm$ SD		P Value
Colorado			
Locked	2.03 $\pm$ 0.38	Colorado vs. Galveston	0.28
Unlocked	1.15 $\pm$ 0.22		
Galveston			
Locked	2.48 $\pm$ 0.57	Locked vs. unlocked	0.003
Unlocked	1.44 $\pm$ 0.35		
Intact spine	0.50 $\pm$ 0.19	Interaction term	0.59

**Table 4. 2-Way ANOVA Results for L5– S1 Torsion Motion (degrees)**

	Mean $\pm$ SD		P Value
Colorado			
Locked	2.02 $\pm$ 0.81	Colorado vs. Galveston	0.11
Unlocked	2.36 $\pm$ 1.01		
Galveston			
Locked	1.23 $\pm$ 0.66	Locked vs. unlocked	0.04
Unlocked	1.36 $\pm$ 0.66		
Intact spine	3.20 $\pm$ 1.64	Interaction term	0.57

designed by Daniel Chopin addresses some of these issues.<sup>15</sup> Multiple points of fixation distribute the loads, and threaded screws are able to resist pullout along their axes to a higher degree than a smooth segment of rod within the ilium. In the comparison performed, there was no statistical difference in overall construct stiffness for both flexion-extension and torsional testing. However, while specifically measuring L5–S1 motion, the Colorado II construct did limit flexion-extension motion to a higher extent, perhaps because of the inclusion of the S1 and S2 screws. This result is in contrast to torsional stability in which the Galveston construct tended to be superior, although not statistically.

These findings are consistent with the understanding of the design differences in these systems. The Galveston rod maintains a continuous rigid connection between the spinal segment of the rod and the iliac portion (*i.e.*, it is the same piece of metal). This in contrast to the Colorado II system in which the spinal rod rigidly connects to the plate, but not the iliac and sacral screws. The screws pass through the plate and are used to hold the plate to the sacropelvis. The 3 posterior-to-anteriorly directed screws powerfully resist posterior displacement of the plate (and rod). However, once any screw loosening occurs, motion across the lumbosacral junction is permitted. The benefit of Colorado II instrumentation for limiting L5–S1 flexion-extension motion comes with additional intraoperative technical demands. Sacropelvic anatomy must be well understood to safely place S1 and S2 screws to avoid the spinal canal, lumbar plexus, and iliac vessels.<sup>23–27</sup> A careful clinical evaluation of these methods will be required to determine if the biomechanical advantage of Colorado II sacropelvic fixation (reduced L5–S1 flexion-extension motion) translates to an improved clinical outcome in these patients with neuromuscular spinal deformity.

One of the more important findings of this study may be the identification of improved Luque-Galveston construct stability of the lumbar spine by simply adding a pair of pedicle screws to the sublaminar wire construct. Theoretically, this procedure fixes the length of the construct as well as adds an undoubtedly more secure point of fixation, together increasing overall stiffness by approximately 50%. The increasing use of pedicle fixation in idiopathic scoliosis has been creeping into neuromuscular constructs as well. These findings support the continued introduction of such fixation into what has traditionally been an exclusively sublaminar wire construct.

## ■ Conclusion

The 2 methods of sacropelvic fixation provided similar construct stiffness, although the Colorado II method had less L5–S1 motion in flexion-extension testing, and the Galveston construct tended to be stiffer in torsional loading. The addition of a pair of L1 pedicle screws increased

the construct stiffness (flexion-extension and torsion) for both constructs by approximately 50%.

## ■ Key Points

- Overall construct stiffness in flexion, extension, and torsional loading was similar between the 2 methods of pelvic fixation.
- Flexion, extension, and L5–S1 motion was smaller with the Colorado sacropelvic fixation.
- The addition of a single pair of L1 pedicle screws to the otherwise Luque wire construct significantly reduced construct stiffness as well as torsional motion at L5–S1.

## References

1. Luque ER. The anatomic basis and development of segmental spinal instrumentation. *Spine* 1982;7:256–9.
2. Luque ER. Segmental spinal instrumentation for correction of scoliosis. *Clin Orthop Relat Res* 1982;163:192–8.
3. Allen BL Jr, Ferguson RL. The Galveston technique for L rod instrumentation of the scoliotic spine. *Spine* 1982;7:276–84.
4. Allen BL Jr, Ferguson RL. The Galveston technique of pelvic fixation with L-rod instrumentation of the spine. *Spine* 1984;9:388–94.
5. Allen BL Jr, Ferguson RL. The Galveston experience with L-rod instrumentation for adolescent idiopathic scoliosis. *Clin Orthop Relat Res* 1988;229:59–69.
6. Bell D, Moseley C, Koreska J. Unit rod segmental spinal instrumentation in the management of patients with progressive neuromuscular spinal deformity. *Spine* 1989;14:1301–7.
7. Maloney WJ, Rinsky LA, Gamble JG. Simultaneous correction of pelvic obliquity, frontal plane, and sagittal plane deformities in neuromuscular scoliosis using a unit rod with segmental sublaminar wires: a preliminary report. *J Pediatr Orthop* 1990;10:742–9.
8. Bulman WA, Dormans JP, Ecker ML, et al. Posterior spinal fusion for scoliosis in patients with cerebral palsy: A comparison of Luque rod and Unit Rod instrumentation. *J Pediatr Orthop* 1996;16:314–23.
9. Dias RC, Miller F, Dabney K, et al. Surgical correction of spinal deformity using a unit rod in children with cerebral palsy. *J Pediatr Orthop* 1996;16:734–40.
10. Rinsky LA. Surgery of spinal deformity in cerebral palsy. Twelve years in the evolution of scoliosis management. *Clin Orthop Relat Res* 1990;253:100–9.
11. Smucker JD, Miller F. Crankshaft effect after posterior spinal fusion and unit rod instrumentation in children with cerebral palsy. *J Pediatr Orthop* 2001;21:108–12.
12. Johnston CE 2nd, Ashman RB, Sherman MC, et al. Mechanical consequences of rod contouring and residual scoliosis in sublaminar segmental instrumentation. *J Orthop Res* 1987;5:206–16.
13. Sink EL, Newton PO, Mubarak SJ, et al. Maintenance of sagittal plane alignment after surgical correction of spinal deformity in patients with cerebral palsy. *Spine* 2003;28:1396–403.
14. Gau YL, Lonstein JE, Winter RB, et al. Luque-Galveston procedure for correction and stabilization of neuromuscular scoliosis and pelvic obliquity: A review of 68 patients. *J Spinal Disord* 1991;4:399–410.
15. Chopin D. A new device for pelvic fixation for spinal surgery: The sacral block. Paper presented at: Eighth Proceeding of the International Congress of Cotrel-Dubousset Instrumentation; 1991; Minneapolis, MN.
16. McCord DH, Cunningham BW, Shono Y, et al. Biomechanical analysis of lumbosacral fixation. *Spine* 1992;17:S235–43.
17. Glazer PA, Colliou O, Klisch SM, et al. Biomechanical analysis of multilevel fixation methods in the lumbar spine. *Spine* 1997;22:171–82.
18. Leong JC, Lu WW, Zheng Y, et al. Comparison of the strengths of lumbosacral fixation achieved with techniques using one and two triangulated sacral screws. *Spine* 1998;23:2289–94.
19. Kornblatt MD, Casey MP, Jacobs RR. Internal fixation in lumbosacral spine fusion. A biomechanical and clinical study. *Clin Orthop Relat Res* 1986;203:141–50.
20. Guyer DW, Yuan HA, Werner FW, et al. Biomechanical comparison of seven internal fixation devices for the lumbosacral junction. *Spine* 1987;12:569–73.
21. Camp JF, Caudle R, Ashmun RD, et al. Immediate complications of Cotrel-

- Dubousset instrumentation to the sacro-pelvis. A clinical and biomechanical study. *Spine* 1990;15:932–41.
22. Miladi LT, Ghanem IB, Draoui MM, et al. Iliosacral screw fixation for pelvic obliquity in neuromuscular scoliosis. A long-term follow-up study. *Spine* 1997;22:1722–9.
23. Esses SI, Botsford DJ, Huler RJ, et al. Surgical anatomy of the sacrum. A guide for rational screw fixation. *Spine* 1991;16:S283–8.
24. Mirkovic S, Abitbol JJ, Steinman J, et al. Anatomic consideration for sacral screw placement. *Spine* 1991;16:S289–94.
25. Asher MA, Strippgen WE. Anthropometric studies of the human sacrum relating to dorsal transsacral implant designs. *Clin Orthop Relat Res* 1986; 203:58–62.
26. Smith SA, Abitbol JJ, Carlson GD, et al. The effects of depth of penetration, screw orientation, and bone density on sacral screw fixation. *Spine* 1993;18: 1006–10.
27. Carlson GD, Abitbol JJ, Anderson DR, et al. Screw fixation in the human sacrum. An in vitro study of the biomechanics of fixation. *Spine* 1992;17: S196–203.



# Early Failures Following Cervical Corpectomy Reconstruction With Titanium Mesh Cages and Anterior Plating

Michael D. Daubs, MD, FACS

## **Study Design.** Retrospective case series.

**Objective.** To evaluate the use of titanium mesh cages in the reconstruction of the cervical spine following corpectomy.

**Summary of Background Data.** Previous studies have shown high successful fusion rates and low failure rates with this technique. Similar reconstruction techniques using anterior strut bone grafting and anterior plating have shown higher failure rates following multilevel corpectomies.

**Methods.** A retrospective review was performed of 23 consecutive patients who underwent anterior cervical corpectomy reconstructed with a titanium mesh cage, local autograft, and fixed anterior plating. Medical records and radiographs were reviewed. Average follow-up was 28 months.

**Results.** Seven patients (30%) had reconstruction failures. There was 1 failure (6%) in the 1-level corpectomy group and 6 (75%) in the multilevel corpectomy group. All failures occurred before 12 weeks after surgery. The remaining patients had successful fusion (70%).

**Conclusion.** There is a high early failure rate (75%) with the use of a titanium mesh cage and fixed anterior plating for reconstruction of multilevel corpectomies. Posterior fusion and instrumentation should be considered when using this technique for multilevel reconstructions.

**Key words:** cervical, corpectomy, multilevel reconstruction, titanium mesh cage, anterior cervical plating, failure. *Spine* 2005;30:1402–1406

The use of titanium mesh cages in the reconstruction of the cervical spine following corpectomy has been reported.<sup>1–4</sup> The potential advantages were immediate anterior column stability, avoidance of bone graft site morbidity, good biocompatibility, and reduced instrumentation morbidity.<sup>1</sup> The reported successful fusion rate ranged from 95% to 100%, and the implant complication rate ranged from 6% to 28%.<sup>1,3,4</sup>

Early construct failures following anterior cervical corpectomy and fusion using autograft and allograft struts and anterior plating have also been reported.<sup>5–9</sup>

The early failure rate with this technique ranged from 6% to 9%<sup>5,6</sup> in 2-level corpectomies and up to 71%<sup>5</sup> in 3 levels. In this study, we reviewed our experience with 1, 2, and 3-level cervical corpectomies reconstructed with a titanium mesh cage, local autograft, and anterior plating in 27 consecutive patients.

## ■ Materials and Methods

A retrospective analysis was performed on 27 consecutive patients who underwent anterior cervical corpectomy and fusion with the use of a titanium mesh cage and anterior cervical plating using a fixed cervical plate. One surgeon (M.D.D.) performed all surgical procedures. Four patients were lost to follow-up. The 4 patients who were lost to follow-up underwent a 1-level corpectomy reconstruction performed for a traumatic cervical burst fracture. Inclusion criteria included cervical stenosis with myelopathy or radiculopathy and cervical stenosis secondary to traumatic burst fracture. Exclusion criteria included tumor and infection.

There were 13 males and 10 females. Mean age was 61 years (range 32–84). The diagnosis was cervical stenosis with myelopathy in 22 patients and cervical burst fracture in 1. Patients were observed for an average of 28 months (range 12–77). Harms cages (Depuy, Raynham, MA) were used in 22 patients and Syn-Mesh (Synthes, Paoli, PA) in 1. All cages were packed with autograft from the resected vertebrae. All cases were stabilized with a fixed anterior cervical plate. A Synthes CSLP plate (Synthes, Paoli, PA) was used in 22 patients and a Depuy Peak plate (Depuy, Raynham, MA) in 1.

A complete corpectomy was performed using a standard channel technique in all patients. The anterior two-thirds of the vertebral body was removed with rongeurs, and the posterior one-third was removed using a high-speed burr. The posterior longitudinal ligament (PLL) was completely removed in all cases. Fifteen patients underwent a 1-level corpectomy, 6 a 2-level, and 2 underwent a 3-level corpectomy. All patients wore an Aspen collar following surgery, except one patient, who was placed in a halo. Complications of construct failure with cage subsidence, and/or cage and plate extrusion were studied.

## ■ Results

Seven patients (30%) had catastrophic failure of fixation with cage subsidence and distal plate extrusion (Table 1). Of the 7 failures, 5 occurred within 6 weeks postoperatively, and all occurred before 12 weeks. There was 1 failure (1 of 15, 6%) in the 1-level corpectomy group, 4 (4 of 6, 67%) in the 2-level group, and 2 (2 of 2, 100%) in the 3-level group. One patient in the 3-level corpectomy group had failure despite postoperative halo immobilization. All of the failures occurred by the same mech-

From the Department of Surgery, University of Nevada School of Medicine, Orthopaedic Specialists, Las Vegas, NV.

Acknowledgment date: April 21, 2004. First revision date: July 13, 2004. Acceptance date: July 16, 2004.

The device(s)/drug(s) that is/are the subject of this manuscript is/are not FDA-approved for this indication and is/are not commercially available in the United States.

No funds were received in support of this work. No benefits in any form have been or will be received from a commercial party related directly or indirectly to the subject of this manuscript.

Address correspondence and reprint requests to Michael D. Daubs, MD, FACS, Department of Surgery, University of Nevada School of Medicine, Orthopaedic Specialists, 701 South Tonopah Drive, Las Vegas, NV 89106; E-mail: Mdd63@aol.com

**Table 1. Patient Data of Failed Reconstructions**

Age (yrs)/Sex	Diagnosis	Anterior Plate/Cage	No. of Corpectomy Levels (fusion levels)	Time of Failure Postoperatively (wks)
62/F	CSM	Depuy/Harms	2 (C4–C7)	10
56/F	CSM	Synthes/Harms	2 (C4–C7)	8
68/F	CSM	Synthes/Harms	3 (C3–C7)	4
52/M	CSM	Synthes/Harms	2 (C4–C7)	2
62/F	CSM	Synthes/Harms	1 (C4–C6)	3
81/M	CSM	Synthes/Harms	3 (C2–C6)	5
74/M	CSM	Synthes/Harms	2 (C3–C6)	7

CSM = cervical stenosis with myelopathy.

anism of cage subsidence into the caudal vertebral body, followed by plate or both cage and plate extrusion (Figure 1). Of the failures, 4 were fused distally to C-7, and 3 were fused to C-6 during the original procedure. One patient underwent complete corpectomy at C-5 with cage placement from C4–C6 and an anterior cervical discectomy and fusion at the C6–C7 level with fibular allograft. The cage subsided into the cephalad C-4 vertebral body (Figure 2). This resulted in a kyphotic deformity, but the construct remained intact and was not revised. Despite the construct failures in this series, no patients worsened neurologically. There were no esophageal, tracheal, or vascular injuries resulting from the

extruded devices. There were no late failures beyond the 12-week follow-up. Successful fusion was obtained in all patients who did not have early construct failure (70%).

Other complications included 1 patient with a postoperative hematoma requiring surgical evacuation, 2 with severe dysphagia, 1 requiring the temporary placement of a gastric feeding tube, and 1 with a C-5 nerve palsy that fully recovered at 6 months after surgery. Revision surgery was performed on 6 of the 7 failures. Despite distal plate dislodgement, one patient, decided against revision surgery because of severe medical problems. The cage was removed in all revised cases and replaced with a fibular strut allograft. Because of the com-

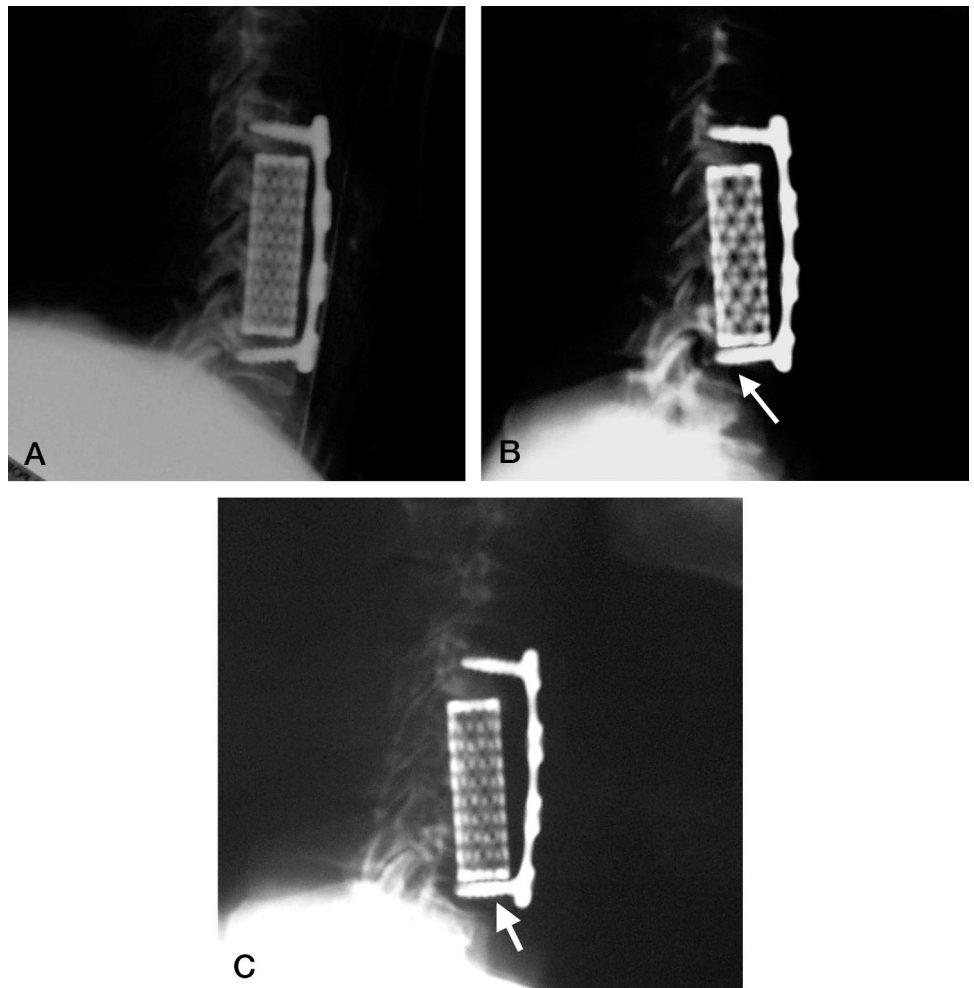


Figure 1. A 62-year-old female who underwent a 2-level corpectomy of C5 and C6 with fusion from C4–C7. **A**, Postoperative lateral radiograph at 2 weeks. **B**, Postoperative lateral radiograph at 6 weeks showing subsidence of the cage into the C-7 vertebral body. **C**, Postoperative lateral radiograph at 10 weeks showing complete subsidence of the cage through the endplate of C-7 into the disc space of C7–T1.

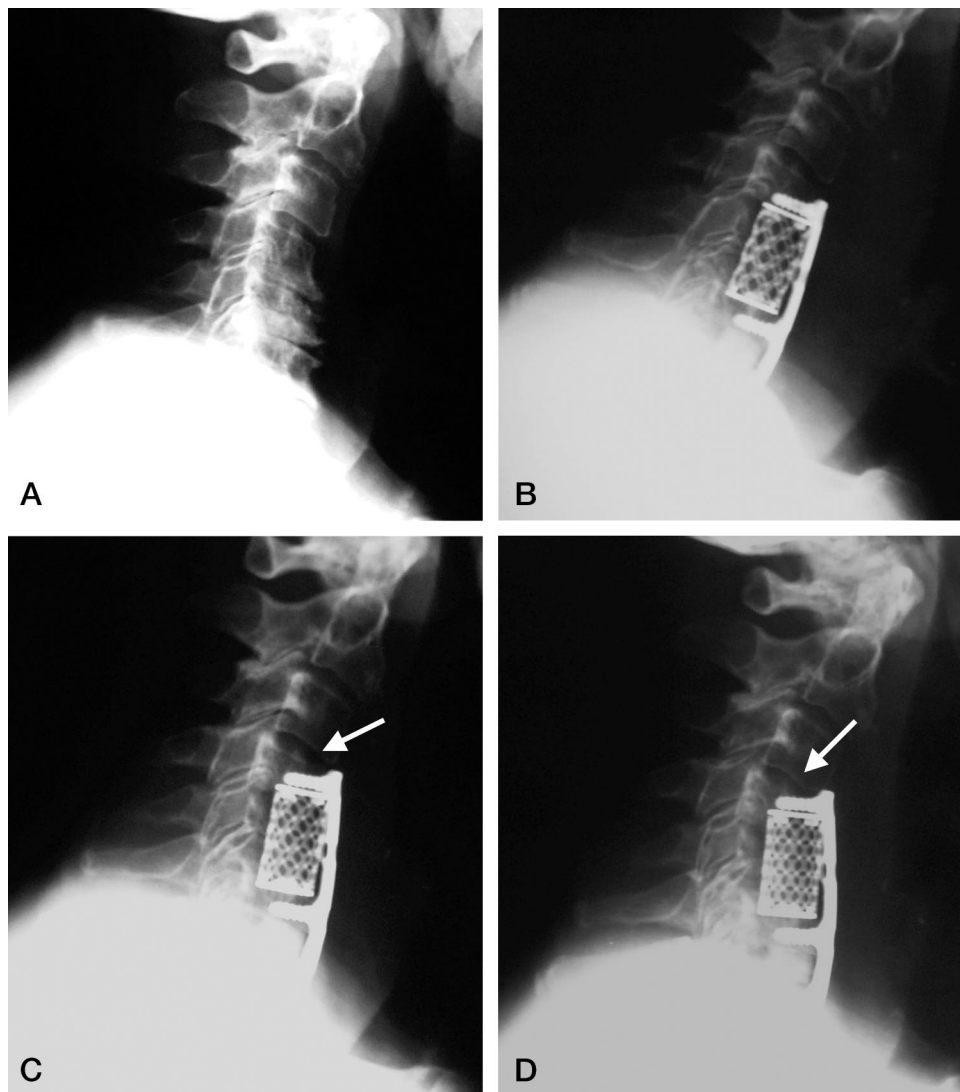


Figure 2. A 70-year-old male with cervical spondylosis and myelopathy. **A**, Preoperative lateral radiograph showing spondylosis and severe disc space collapse with kyphosis. **B**, Postoperative lateral radiograph at 2 weeks showing the Syn-Mesh cage placed from C4–C6 at the level of the C5 corpectomy, and a fibular allograft placed at the level of C6–C7 where a discectomy was performed. The Synthes CSLP plate was placed from C4–C7, with screws inserted distally at C6 and C7. **C**, Postoperative lateral radiograph at 6 weeks showing superior subsidence of the cage (arrow) into the C4 vertebral body. **D**, Postoperative lateral radiograph at 9 weeks showing further subsidence of the cage (arrow) and screws into C4, and further collapse into kyphosis.

plete destruction of the caudal vertebral body by the subsiding cage, all patients required an additional 1-level corpectomy and an extension of the fusion 1-level distally. All revisions included a posterior approach with supplemental posterior instrumentation and fusion using iliac crest bone graft. There were no implant failures following the revision surgery at last follow-up, and all patients had successful fusion. One patient died 3 months following revision surgery of a myocardial infarct.

## Discussion

Previous reports on the use of titanium mesh cages and anterior plating after cervical corpectomies have been favorable, with successful fusion rates ranging from 95% to 100% and implant complication rates 6% to 28%.<sup>1–4</sup> The initial report of this technique by Majd *et al*<sup>1</sup> showed a 97% successful fusion rate and a clinical outcome of excellent or good in 80% of their patients. They reported one cage extrusion (2.9%).

Thalgott *et al*<sup>4</sup> had a 100% fusion rate and an excellent or good outcome in 81% of their patients, with only one patient having significant subsidence of the cage and

screw back out. The reported plate screw complication rate was 15% in that series, with no reported cage extrusions. One patient required revision surgery for failure of plate fixation, but the cage was not revised. Hee *et al*<sup>3</sup> showed specifically on the complications of this technique in multilevel corpectomies. Their overall complication rate was 33%. Of their 21 patients, 6 (28%) had cage, plate, or screw complications. Three (14%) required revision surgery. Of the 3 patients who had significant subsidence, 2 had osteopenia.

Our failure rate was 6% (1 of 15) in the single level corpectomy group and 75% (6 of 8) in the multilevel corpectomy group. The single level corpectomy group in this series had a similar failure rate to that of strut grafting, but the multilevel group had a higher rate of failure.<sup>5,6,8,9</sup> The higher failure rate in this study may be a result of increased patient age and osteopenia, the inherent biomechanical instability of long segment anterior cervical fixation, and the unique properties of the titanium mesh cage.

The patients in this study were typical of the older patient population with cervical spondylotic myelopathy, in which a wide and complete corpectomy is re-



quired for adequate spinal cord decompression. Overall average age in this study was 61 years, and average age in the failed group was 65 years. This study group differs from the population reported by Thalgott *et al*,<sup>4</sup> in which average age was 51, no patients had myelopathy, and the corpectomy was performed primarily as a method to reduce the number of bone-graft surface areas for fusion. This was also the case in the original study by Majd *et al*,<sup>1</sup> in which only 14 of their 34 patients had myelopathy and underwent a complete corpectomy with removal of the PLL for decompression. In the remainder of the cases, the vertebra was only partially removed to accommodate cage placement. In this series, a complete corpectomy with removal of the PLL and anulus structures was performed in all patients. The more aggressive bony decompression and complete resection of the PLL may have increased the instability of the cervical spine<sup>10,11</sup> and, as a result, reduced the stability of reconstruction. In the study by Hee *et al*,<sup>3</sup> average age was 57 years, and they reported osteopenia in 2 of the 3 failures. Of the 7 patients who had failure in this study, 6 were osteopenic.

Lim *et al*<sup>11</sup> studied the association of endplate condition and bone mineral density on the compressive strength of the graft endplate interface in the cervical spine. Load-to-failure decreased with decreasing bone mineral density and progressive removal of the endplate. Although meticulous effort was made to preserve the endplate in our cases, the sharp teeth of the cage often penetrated the endplate at insertion. Because this may increase the resistance to initial anterior dislodgement and add to the initial stability as noted by Hee *et al*,<sup>3</sup> it also may lead to increase subsidence and failure by cutting through the endplate. The small footprint and surface area of contact of the mesh cage with the vertebral endplate may also increase subsidence. End-caps for the mesh cages were later introduced to avoid the problem of sharp edges and to increase the surface area of contact. End-caps were used in 3 of the 7 failures in this series (Figure 1). In partially decorticated or osteoporotic vertebral endplates, the increased surface area of contact may not be enough to resist compression and subsidence. Although the end-cap does increase the surface area of contact of the cage with the vertebral body endplate, it decreases the surface area of contact with the bone graft, which may affect the fusion rate.

Subsidence can also occur with the use of a cortical fibular allograft,<sup>5-9</sup> especially in partially decorticated or osteoporotic vertebral endplates. Whitecloud and LaRocca<sup>12</sup> described a technique aimed at better resisting the compressive forces on the endplates by forming a notch at each end of the cortical allograft, and locking it onto the more compressive resistant anterior cortices of the superior and inferior vertebral bodies. To the best of our knowledge, a similar technique has not been reported with the use of titanium mesh cages. However, a larger diameter cage that loads the periphery of the endplates and the outer cortices of the vertebral body could potentially result in less subsidence.

Foley *et al*<sup>13</sup> studied the effects of anterior plating on the stability of the cervical spine after multilevel corpectomies and strut grafting. They concluded that anterior plating increased stiffness, but it also caused loading in extension that might lead to pistoning and failure. In a similar cadaver model, Isomi<sup>14</sup> and Panjabi<sup>15</sup> *et al* evaluated the stabilizing capacity of an anterior plate and strut graft in a 1 and 3-level corpectomy model before and after fatigue loading. They found no difference between the 1 and 3-level models before fatigue, but the stability of the 3-level models significantly decreased after fatigue testing. Fatigue instability likely accounted for the difference in failure rates between the single and multilevel cases in this series. In addition, the titanium mesh cage, with its small footprint and sharp edge, may be more susceptible to fatigue failure and subsidence through a pistoning mechanism.

Although this is a single surgeon case series and a larger multicenter cohort study is needed to better evaluate this technique, the results suggest that titanium mesh cage and anterior plating have a higher failure rate than that reported with strut grafting and anterior plating in multilevel corpectomy reconstructions. The unique properties of the mesh cage, osteopenia, and the known biomechanical shortfalls of anterior fixation alone following multilevel cervical corpectomies may have all played a role in the high failure rate in this study.

## ■ Conclusion

Although previous investigators have touted the benefits of the titanium mesh cage in reconstruction of the cervical spine following cervical corpectomy, its use in multilevel corpectomy reconstructions should be cautious. The failure rate in this case series was higher than that reported with the use of strut bone grafting and anterior plating in multilevel corpectomy reconstructions. The patients in this series are representative of the patient population with cervical spondylotic myelopathy, in which an extensive anterior decompression is necessary, and poor bone quality is a factor. The titanium mesh cage, with its sharp edges and small surface area of contact, may be more susceptible to pistoning, subsidence, and eventual fatigue failure, especially in those patients with osteopenia. If this technique is used in reconstruction of multilevel cervical corpectomies, posterior stabilization and instrumentation should be considered.

## ■ Key Points

- The success rate is satisfactory with a one-level corpectomy.
- The failure rate was high (75%) in multilevel corpectomies.
- Caution should be used with this technique in multilevel reconstructions and in patients with osteopenia.

## References

1. Madj ME, Vadhva M, Holt RT. Anterior cervical reconstruction using titanium cages with anterior plating. *Spine* 1999;24:1604–10.
2. Das K, Couldwell WT, Sava G, et al. Use of cylindrical titanium mesh and locking plates in anterior cervical fusion. Technical note. *J Neurosurg Spine* 2001;94:174–8.
3. Hee HT, Madj ME, Holt RT, et al. Complications of multilevel cervical corpectomies and reconstruction with titanium cages and anterior plating. *J Spinal Disord* 2003;16:1–8.
4. Thalgott JS, Xiongsheng C, Giuffre JM. Single stage anterior cervical reconstruction with titanium mesh cages, local bone graft, and anterior plating. *Spine J* 2003;3:294–300.
5. Sasso RC, Ruggiero RA, Reilly TM, et al. Early reconstruction failures after multilevel cervical corpectomy. *Spine* 2003;28:140–2.
6. Vaccaro AR, Falatyn SP, Scuderi GJ, et al. Early failure of long segment anterior cervical plate fixation. *J Spinal Disord* 1998;11:410–5.
7. Zeidman SM, Ducker TB, Raycroft J. Trends and complications in cervical spine surgery. *J Spinal Disord* 1997;10:523–6.
8. Mayr MT, Subach BR, Comey CH, et al. Cervical spinal stenosis: Outcome after anterior corpectomy, allograft reconstruction, and instrumentation. *J Neurosurg* 2002;96:10–6.
9. Macdonald RL, Fehlings MG, Tator CH, et al. Multilevel anterior cervical corpectomy and fibular allograft fusion for cervical myelopathy. *J Neurosurg* 1997;86:990–7.
10. McAfee PC, Cunningham B, Dmitriev A, et al. Cervical disc replacement-porous coated motion prosthesis: A comparative biomechanical analysis showing the key role of the posterior longitudinal ligament. *Spine* 2003;28:S176–85.
11. Lim TH, Kwon H, Jeon CH, et al. Effect of endplate conditions and bone mineral density on the compressive strength of the graft-endplate interface in anterior cervical spine fusion. *Spine* 2001;26:951–6.
12. Whitecloud T, LaRocca H. Fibular strut graft in reconstructive surgery of the cervical spine. *Spine* 1976;1:33.
13. Foley KT, DiAngelo DJ, Rampersaud YR, et al. The in-vitro effects of instrumentation on multilevel cervical strut-graft mechanics. *Spine* 1999;24:2366–76.
14. Isomi T, Panjabi MM, Wang JL, et al. Stabilizing potential of anterior cervical plates in multilevel corpectomies. *Spine* 1999;24:2219–23.
15. Panjabi MM, Isomi T, Wang JL. Loosening at the screw-vertebra junction in multilevel anterior cervical plate constructs. *Spine* 1999;24:2383–8.

# Correlation Between Range of Motion and Outcome After Lumbar Total Disc Replacement: 8.6-Year Follow-up

Russel C. Huang, MD,\* Federico P. Girardi, MD,\* Frank P. Cammisa, Jr., MD,\*  
Moe R. Lim, MD,\* Patrick Tropiano, MD,† and Thierry Marnay, MD‡

**Study Design.** Retrospective radiographic and chart review.

**Objective.** To examine the relationship between lumbar total disc replacement (TDR) range of motion (ROM) and clinical outcome at 8.6-year follow-up.

**Summary of Background Data.** There are no studies on the relationship between TDR motion and clinical outcomes.

**Methods.** We reviewed 38 patients who underwent 1 or 2-level TDR implantation with 51 TDR. Flexion-extension ROM was measured on lateral radiographs. Clinical outcomes were measured at 8.6 years by modified Stauffer-Coventry scores, Oswestry Disability Questionnaires (ODQ), and subjective ratings of back pain, leg pain, and disability. Spearman rank correlation coefficient was used to determine if ROM was correlated with clinical outcome. Patients were divided into 2 groups by motion ( $\leq 5^\circ$  and  $> 5^\circ$ ). Statistical differences in outcome were sought.

**Results.** Spearman rank correlation coefficient revealed weak-to-moderate but statistically significant associations between ROM and outcome for postoperative back pain ( $r = -0.35$ ,  $P = 0.034$ ), ODQ ( $r = -0.33$ ,  $P = 0.046$ ), and modified Stauffer-Coventry scores ( $r = 0.42$ ,  $P = 0.0095$ ). Patients with motion of  $> 5^\circ$  had superior outcomes in ODQ (mean difference 12.6 points,  $P = 0.026$ ) and Stauffer-Coventry scores (mean difference 2.2 points,  $P = 0.015$ ).

**Conclusions.** The radiographic ROM at 8.6-year follow-up was positively correlated with several outcomes measures. Patients with motion  $> 5^\circ$  had clinically modest but statistically better outcomes in ODQ and modified

Stauffer-Coventry scores. Longer follow-ups will be necessary to measure fully the impact of TDR ROM on outcome.

**Key words:** total disc replacement, arthroplasty, lumbar, spine, range of motion, outcomes, correlation. **Spine** 2005;30:1407–1411

One of the fundamental theoretical advantages of total disc replacement (TDR) over fusion is the preservation of segmental motion. Increased stresses on mobile segments adjacent to fusions accelerate adjacent level degeneration, which may present as combinations of disc degeneration, facet arthrosis, instability, and spinal stenosis.<sup>1–5</sup> Adjacent level degeneration after fusion is a significant clinical problem that frequently requires revision surgery.<sup>1,3,5,6</sup>

Lumbar TDR has been performed on more than 5000 patients worldwide. Multiple investigators<sup>7–10</sup> have reported the range of motion (ROM) of these implants and clinical outcomes, but none have carefully examined the relationship between ROM and outcome. If the ROM of TDR implants is not statistically correlated to outcome, then the justification for its use would seem tenuous. In an investigation of the same patient cohort analyzed in the current study, Huang *et al*<sup>7</sup> showed that there is a statistically significant association between radiographic adjacent level degeneration and low TDR ROM at a longer than 8-year follow-up, suggesting that TDR ROM may affect radiographic outcome. The establishment of a clear relationship between TDR ROM and clinical outcome would provide support for the theoretical basis of TDR (*i.e.*, that the preservation of segmental motion may improve outcomes). The purpose of this article is to examine statistically the relationship between lumbar TDR ROM and clinical outcome at 8.6-year follow-up.

## ■ Materials and Methods

**Patients.** Between March 1990 and September 1993, 93 ProDisc® TDRs (Aesclup AG & Co., Tuttlingen, Germany) were implanted in 64 patients by a single surgeon (T.M.). Indication for surgery was disc degeneration with discogenic pain that had failed at least 6 months of nonsurgical treatment. Final clinical and radiographic follow-up was performed between May 1999 and January 2001. Of the 64 patients in the initial cohort, 3 were deceased, and 3 were lost to follow-up. Sixteen patients were excluded from this study because postoperative flexion-extension radiographs were unavailable. The 4 patients who had undergone 3-level implantation were excluded to simplify

From the \*Spine Surgery Service, Hospital for Special Surgery, New York, NY, †Department of Orthopaedic Surgery, Hôpital CHU Nord, Marseille, France, and ‡Department of Orthopaedic Surgery, Clinique du Parc, Castelnau-le-Lez, France.

Acknowledgment date: January 29, 2004. First revision date: May 4, 2004. Second revision date: July 21, 2004. Acceptance date: August 2, 2004.

The device(s)/drug(s) that is/are the subject of this manuscript is/are not FDA-approved for this indication and is/are not commercially available in the United States. The device is no longer in use, is no longer manufactured, and was used only in France.

Corporate/Industry funds were received to support this work. Although one or more of the author(s) has/have received or will receive benefits for personal or professional use from a commercial party related directly or indirectly to the subject of this manuscript, benefits will be directed solely to a research fund, foundation, educational institution, or other nonprofit organization which the author(s) has/have been associated. One or more of the author(s) has/have received or will receive benefits for personal or professional use from a commercial party related directly or indirectly to the subject of this manuscript: e.g., royalties, stocks, stock options, decision making position.

Address correspondence and reprint requests to Russel C. Huang, MD, Hospital for Special Surgery, 535 East 70<sup>th</sup> Street, New York, NY 10021; E-mail: russelhuang@yahoo.com



**Table 1. Distribution of Levels Implanted**

Levels Implanted (No. level)	No. of Patients
L3–L4 (1)	1
L4–L5 (1)	16
L5–S1 (1)	8
L3–L5 (2)	2
L4–S1 (2)	11

data analysis, and because the “middle” level in a 3-level implantation has no native junctional discs. Furthermore, the patient who undergoes 3-level implantation is a rare clinical entity. Included in this study were 38 patients who had undergone single or 2-level implantation, with a total of 51 ProDisc® TDR.

There were 22 male and 16 female patients. At surgery, mean age was  $44.9 \pm 7.7$  years (range 25–65), and mean weight was  $72.0 \pm 14.2$  kg (range 48–102). Mean follow-up was  $8.6 \pm 0.97$  years (range 6.9–10.7). Nineteen (50%) patients had a history of prior spine surgery. A total of 36 prior procedures were performed on these 19 patients, including 22 discectomies, 4 decompressive laminectomies, and 10 other procedures, including thermocoagulation and chemonucleolysis. Levels implanted are summarized in Table 1.

**Measurement of Flexion-Extension ROM.** Postoperative active flexion-extension lateral radiographs were used to determine postoperative ROM. Flexion-extension ROM was measured by the Cobb method, as previously described, at every instrumented level by a single observer (R.C.H.) who was not involved in patient selection, surgery, or follow-up.<sup>7</sup> The intraobserver and interobserver reliability of the Cobb method is reported to be relatively high. Using the Cobb method to measure segmental lumbar lordosis, mean absolute difference between measurements performed by 3 independent observers (mean interclass correlation coefficient 0.87, mean intra-class correlation coefficient 0.93) was reported as  $1.4^\circ$  from L1–S1.<sup>11</sup> Another study specifically addressing accuracy in measurement of ROM after lumbar TDR showed that mean absolute difference between repeated measurements was  $1.6^\circ$  (intraobserver) and  $1.8^\circ$  (interobserver).<sup>11a</sup>

**Clinical Outcomes Measurement.** The patients were evaluated both preoperatively and postoperatively by one of the authors (P.T.) and by a research assistant. These patient evaluators were not involved in patient selection, surgery, or postoperative care. Five outcomes instruments were used to evaluate patients preoperatively and postoperatively. A 20-point

modified Stauffer-Coventry score<sup>10,12</sup> was calculated both preoperatively and postoperatively. The score incorporates pain, neurologic deficit, need for medication, disability, and psychiatric status (Table 2). Possible scores range from 0 (worst) to 20 (best). Patients scored 3 separate subjective parameters on quantitative scales preoperatively and postoperatively:

1. Low back pain was graded as severe (3), moderate (2), mild (1), or absent (0).
2. Leg (radicular) pain was graded as severe (3), moderate (2), mild (1), or absent (0).
3. Disability was quantified by eliciting the ability to perform employment and activities of daily living as normal (0), slightly limited (1), significantly impaired (2), or severely limited/impossible (3).

The Oswestry Disability Questionnaire (ODQ) was only administered after surgery because it was not in use at the senior author's institution at the initial surgeries.

**Statistical Analysis.** Because our sample size was relatively small, patients who had undergone 1 and 2-level TDR were analyzed as a single cohort to achieve sufficient statistical power. In patients with 2-level implantation, average ROM was used for purposes of statistical analysis. For example, a patient with TDR at L4–L5 and L5–S1, with motion of  $3^\circ$  at L4–L5 and  $7^\circ$  at L5–S1 would be entered into the analysis as having  $5^\circ$  ROM. Averaging was performed so that each patient would have one value for ROM and one value for each outcome measure, enabling statistical analysis. We believe that averaging is a valid approach to the analysis of these data because angular deflection is additive across levels. In other words, if the sum total of angular motion across 2 disc spaces is  $10^\circ$ , the total angulation remains  $10^\circ$  regardless of how many degrees are contributed by each disc space.

To determine if a statistical correlation exists between ROM and outcome, Spearman rank correlation coefficient was performed for mean TDR ROM *versus* postoperative back pain, leg pain, Oswestry Disability scores, and modified Stauffer-Coventry scores. All outcomes scores were considered to be ordinal, nonparametric data. Spearman rank (*r*) with 95% confidence intervals and *P* values are reported.

To assess the clinical significance of TDR ROM, the patients were divided into those with  $\leq 5^\circ$  ROM and those with  $> 5^\circ$  motion. We selected  $5^\circ$  as a dividing line based on its use as a fusion threshold in Food and Drug Administration trials.<sup>13</sup> Mean postoperative clinical outcomes scores for each group were evaluated for statistical differences with Mann-Whitney testing. Preoperative data were evaluated by the Student *t*

**Table 2. Modified Stauffer-Coventry Score**

Parameter	Grade					Total
	0	1	2	3	4	
Low back pain	Permanent	Frequent	Moderate	None		3
Radicular pain	Permanent	Frequent	Effort	None		3
Neurologic deficit	Major		Moderate		None	4
Medication	Major	Moderate	None			2
Day living activities	Impossible		Normal			2
Work status postoperatively	No work	Frequently stopped	Change	Same work >6 mos. Change <3 mos	Normal	4
Psychiatric status	Preoperative	Secondary to pathology	None			2
Total		20				

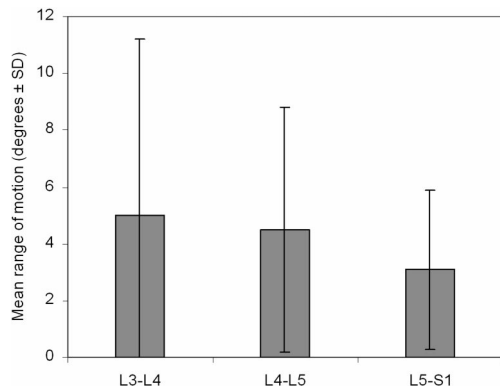


Figure 1. Mean TDR ROM (degrees  $\pm$  standard deviation [SD]) by level.

(parametric data), Mann-Whitney (ordinal data), and Fisher exact (nominal data) tests to ensure that the 2 groups of patients were similar before surgery. All  $P$  values reported are 2-tailed. The alpha level of statistical significance was 0.05. All numerical values are reported as a number  $\pm$  standard deviation (range) where applicable. All data analysis was performed using Microsoft® Excel 2002 (Microsoft Inc., Redmond, WA) and InStat 3.05 (GraphPad Software Inc., San Diego, CA).

## Results

Mean TDR ROM in this cohort of patients was  $4.0^\circ \pm 3.9^\circ$  (range  $0^\circ$  to  $18^\circ$ ). Mean ROM by level is summarized in Figure 1. Spearman correlation revealed weak-to-moderate but statistically significant associations between ROM and outcome for postoperative back pain, Oswestry disability scores, and modified Stauffer-Coventry scores. There was no statistically significant correlation between ROM and postoperative leg pain scores (Table 3).

When divided into subgroups based on ROM, 28 patients had ROM  $\leq 5^\circ$ , and 10 patients had ROM  $> 5^\circ$ . Mean ROM in the low ROM group was  $2.2^\circ \pm 1.5^\circ$  (range  $0^\circ$  to  $5^\circ$ ). In the high ROM group, mean motion was  $9.8^\circ \pm 3.9^\circ$  (range  $6^\circ$  to  $18^\circ$ ). Analysis of data showed that before surgery, both groups were similar regarding age, sex, weight, history of previous surgery, back and leg pain, disability, and modified Stauffer-Coventry scores (Table 4). All  $P$  values were  $> 0.05$ .

Analysis of postoperative data showed that there were clinically modest but statistically significant differences in disability (3-point subjective scale and ODQ), and modified Stauffer-Coventry scores between low and high motion groups. There was a trend toward lower low back pain scores in the high ROM group, but it was not

statistically significant ( $P = 0.080$ ). Data are summarized in Table 5.

## Discussion

Preservation of ROM is one of the fundamental theoretical advantages of TDR over arthrodesis. The phenomenon of adjacent level degeneration after fusion is well known and frequently symptomatic. However, despite the fact that more than 5000 patients have been implanted with TDR worldwide, convincing evidence that motion preservation affects outcomes is lacking. Ultimately, randomized prospective trials comparing TDR with fusion and nonsurgical treatment with long-term follow-up will be needed to define the true effectiveness of TDR. Until then, careful analysis of available data from case series is essential to provide preliminary information. Preserved ROM of at least  $2^\circ$  has been documented in 66% of ProDisc® TDR at 8.7-year follow-up.<sup>7</sup> Furthermore, a statistically significant association between low ROM and the development of radiographic adjacent level degeneration has been observed. However, there are no published reports exploring the relationship between TDR ROM and clinical outcome.

The results of the current study reveal a weak-to-moderate (Spearman  $r \approx \pm 0.35$ ) but statistically significant relationship between TDR flexion-extension ROM and clinical outcomes at 8.6-year follow-up using several outcomes instruments. The data further show that patients with low ROM ( $\leq 5^\circ$ ) have slightly inferior outcomes compared to patients with more motion ( $> 5^\circ$ ) regarding Oswestry Disability and modified Stauffer-Coventry scores. Both subgroups of patients were similar before surgery.

Because none of the patients in this study had arthrodesis, it is impossible to make any direct comparisons between fusion and motion preserving treatments. However, the TDR ROM our patients had was scattered across a wide range of values. Patients with solid arthrodeses may be kinematically more similar to those with low TDR ROM than to patients with high ROM. It will be interesting to see if longer term follow-up in prospective randomized trials of fusion *versus* TDR reveals similar trends.

It is not surprising that the magnitude of the correlation between TDR ROM and outcome (Spearman  $r \approx \pm 0.35$ ) is modest. A perfect correlation suggests that 100% of the variance in one variable (*e.g.*, outcome) is accounted for by the variance in another variable (*e.g.*,

Table 3. Spearman Correlation of ROM With Postoperative Outcome for All Patients

	Mean Score	Spearman $r$	95% Confidence Interval	$P$ Value
Back pain	$1.32 \pm 0.81$	-0.35	-0.61 to -0.019	0.034*
Leg pain	$0.66 \pm 0.88$	-0.12	-0.43 to +0.22	0.48
Oswestry Disability Questionnaire	$18.2 \pm 16.0$	-0.33	-0.59 to -0.0033	0.046*
Modified Stauffer-Coventry score	$16.2 \pm 2.6$	+0.42	+0.10 to +0.65	0.0095*

\*Statistically significant.

**Table 4. Comparison of Preoperative Data Among Patients With  $\leq 5^\circ$  and  $>5^\circ$  ROM**

	$\leq 5^\circ$ ROM	$>5^\circ$ ROM	P Value
Age	45.1 $\pm$ 5.5	44.1 $\pm$ 12.2	0.72
Weight	73.7 $\pm$ 15.3	67.9 $\pm$ 11.0	0.29
Males	60.7%	50.0%	0.71
Previous lumbar surgery	46.4%	60.0%	0.71
Back pain	2.8 $\pm$ 0.50	2.6 $\pm$ 0.70	0.56
Leg pain	2.6 $\pm$ 0.88	2.3 $\pm$ 1.3	0.64
Disability	3.1 $\pm$ 0.72	2.5 $\pm$ 0.97	0.12
Modified Stauffer-Coventry score	7.1 $\pm$ 2.9	6.6 $\pm$ 3.9	0.87

ROM). Surgical outcomes are multifactorial. Significant factors include patient variables both related and unrelated to spinal pathology, and technical factors related to surgical technique and implant characteristics. No single variable would be expected to account for a large percentage of variance in outcome.

The ideal amount of motion a TDR should have is unknown. Although it is generally agreed that inadequate motion (*e.g.*, in the case of fusion) accelerates degeneration of adjacent levels, the amount of motion required to avoid junctional degeneration is unknown, but some intriguing data are available. In an investigation of the same patient cohort analyzed in the current study, Huang *et al*<sup>7</sup> found that 24% of patients had radiographic junctional degeneration at 8.7 years postoperatively. Closer analysis of this cohort revealed that no patients with TDR ROM  $\geq 5^\circ$  had adjacent level degeneration, suggesting that  $5^\circ$  of motion may represent a protective threshold against adjacent level degeneration (unpublished data).<sup>13a</sup> Furthermore, 59% of patients in the group without junctional degeneration had motion  $<5^\circ$ , showing that low motion is necessary but not sufficient to cause junctional degeneration at 8.7 years.

In contrast, it is theoretically possible that excessive TDR motion might result in facet arthrosis and hypertrophy, stenosis, and pain. The use of nonfusion technologies to avoid the known problem of adjacent level degeneration may result in a new clinical entity (*i.e.*, same level degeneration). Our data do not suggest that this is a significant clinical problem at 8.6-year follow-up because it appears that it is advantageous to have  $>5^\circ$  of TDR motion. However, an 8.6-year follow-up is only midterm for an arthroplasty, and our average patient was only 54 years old at last follow-up. It is possible that

increased motion may prove detrimental in longer term follow-up if symptomatic same level facet arthrosis or stenosis develops. Fused levels are known to be immune to progression of same level disease over long-term follow-up.

It is currently unknown why some patients have more motion than others after TDR. It is well known that there is wide variability in lumbar flexion-extension ROM between different levels and between different individuals,<sup>14</sup> but, similarly, the cause for this is not understood. Preoperative motion is probably a significant factor in postoperative motion. This has been clearly shown in other types of arthroplasty,<sup>15</sup> but because we lacked preoperative flexion-extension radiographs, evaluation of preoperative motion was impossible in the current study. Technical factors probably also play a role in ROM because the location of a TDR relative to the instantaneous axis of rotation of the motion segment has been shown in clinical and biomechanical studies to affect ROM.<sup>9,16</sup> All TDR implants are not the same, and it has been suggested that implants with different designs display different *in vivo* kinematics.<sup>17</sup> Finally, postoperative protocols have been shown to affect motion, with early mobilization without bracing yielding more motion.<sup>9</sup> None of the patients in this cohort were braced after surgery.

Correlation shows association and not causation. Our data show that there is a statistically significant association between ROM and outcome, but it is impossible to determine whether good outcomes resulted from good ROM. Indeed, it is impossible to establish causation in this particular relationship even with randomized prospective trials. One alternative explanation for the statistical association we observed is that patients with poor outcome (*e.g.*, more back pain) had pain during flexion-extension radiographs and voluntarily restricted their motion. It is also possible that a third variable (*e.g.*, genetic predisposition) affected both clinical outcome and ROM. It is impossible to establish more than association given the available data, but a statistically significant correlation has been shown, and it is certainly plausible that clinical outcomes resulted in part from ROM.

There are many limitations in the current study. It is a retrospective case series using a combination of validated and unvalidated outcomes instruments. The correlation between motion and outcome is an association, and causation cannot be proven. Our findings are specific to the implant studied and cannot necessarily be generalized to all implants. It takes many years for junctional degeneration to compromise the results of fusion, and an 8.6-year follow-up is midterm. Longer follow-ups will be necessary to evaluate the true safety and efficacy of TDR. Finally, because preoperative magnetic resonance imaging was not available for all patients, it was impossible to define precisely the preoperative condition of intervertebral discs that did not undergo TDR. The preoperative status of these discs may have influenced clinical outcome.

**Table 5. Comparison of Postoperative Data Among Patients With  $\leq 5^\circ$  and  $>5^\circ$  ROM**

	$\leq 5^\circ$ ROM	$>5^\circ$ ROM	P Value
Back pain	1.46 $\pm$ 0.79	0.90 $\pm$ 0.74	0.080
Leg pain	0.71 $\pm$ 0.90	0.50 $\pm$ 0.85	0.51
Disability	1.89 $\pm$ 0.74	1.20 $\pm$ 0.42	0.016*
Oswestry Disability Questionnaire	21.6 $\pm$ 16.6	9.0 $\pm$ 10.2	0.026*
Modified Stauffer-Coventry score	15.6 $\pm$ 2.6	17.8 $\pm$ 1.9	0.015*

\*Statistically significant.



## ■ Conclusion

The current study shows that there is a weak-to-moderate but statistically significant correlation between TDR flexion-extension ROM and clinical outcome at 8.6-year follow-up. In addition, statistically significant but clinically modest improvements in outcome were observed in patients with  $>5^\circ$  ROM.

## ■ Key Points

- At 8.6-year follow-up, TDR ROM was positively correlated with better clinical outcomes for back pain, Oswestry Disability Questionnaires, and modified Stauffer-Coventry scores.
- Patients with ROM  $>5^\circ$  had better Oswestry Disability Questionnaires and Stauffer-Coventry scores. Differences were statistically significant and clinically modest.

## Acknowledgment

The authors thank Ms. Nathalie Larive and Mrs. Pascale Munoz for their invaluable assistance.

## References

1. Balderston RA, Albert TJ, McIntosh T, et al. Magnetic resonance imaging analysis of lumbar disc changes below scoliosis fusions. A prospective study. *Spine* 1998;23:54–8.
2. Kumar MN, Jacquot F, Hall H. Long-term follow-up of functional outcomes and radiographic changes at adjacent levels following lumbar spine fusion for degenerative disc disease. *Eur Spine J* 2001;10:309–13.
3. Lee CK. Accelerated degeneration of the segment adjacent to a lumbar fusion. *Spine* 1988;13:375–7.
4. Wada E, Suzuki S, Kanazawa A, et al. Subtotal corpectomy versus laminoplasty for multilevel cervical spondylotic myelopathy: A long-term follow-up study over 10 years. *Spine* 2001;26:1443–7.
5. Gillet P. The fate of the adjacent motion segments after lumbar fusion. *Spine* 2003;28:338–45.
6. Schlegel JD, Smith JA, Schleusener RL. Lumbar motion segment pathology adjacent to thoracolumbar, lumbar, and lumbosacral fusions. *Spine* 1996;21:970–81.
7. Huang RC, Girardi FP, Cammisa FP, et al. Long-term flexion-extension range of motion of the Prodisc total disc replacement. *J Spinal Disord Tech* 2003;16:435–40.
8. Bertagnoli R, Kumar S. Indications for full prosthetic disc arthroplasty: A correlation of clinical outcome against a variety of indications. *Eur Spine J* 2002;11:S131–6.
9. Cinotti G, David T, Postacchini F. Results of disc prosthesis after a minimum follow-up period of 2 years. *Spine* 1996;21:995–1000.
10. Lemaire JP, Skalli W, Lavaste F, et al. Intervertebral disc prosthesis. Results and prospects for the year 2000. *Clin Orthop Relat Res* 1997;337:64–76.
11. Harrison DE, Harrison DD, Cailliet R, et al. Radiographic analysis of lumbar lordosis: Centroid, Cobb, TRALL, and Harrison posterior tangent methods. *Spine* 2001;26:E235–42.
- 11a. Lim MR, Girardi FP, Zhang K, et al. Measurement of total disc replacement radiographic range-of-motion: A comparison of two techniques. *J Spinal Disord Tech*, Accepted for publication, 2005.
12. Lassale B, Deburge A, Benoist M. Long-term results of the surgical treatment of lumbar stenosis [in French]. *Rev Rhum Mal Osteoartic* 1985;52:27–33.
13. Santos ER, Goss DG, Morcom RK, et al. Radiologic assessment of interbody fusion using carbon fiber cages. *Spine* 2003;28:997–1001.
- 13a. Huang RC, Girardi FP, Lim MR, et al. Range of motion and adjacent level degeneration after lumbar total disc replacement. *Spine J*, Accepted for publication, 2005.
14. Hayes MA, Howard TC, Gruel CR, et al. Roentgenographic evaluation of lumbar spine flexion-extension in asymptomatic individuals. *Spine* 1989;14:327–31.
15. Ritter MA, Harty LD, Davis KE, et al. Predicting range of motion after total knee arthroplasty. Clustering, log-linear regression, and regression tree analysis. *J Bone Joint Surg Am* 2003;85-A:1278–85.
16. Dooris AP, Goel VK, Grosland NM, et al. Load-sharing between anterior and posterior elements in a lumbar motion segment implanted with an artificial disc. *Spine* 2001;26:E122–9.
17. Huang RC, Girardi FP, Cammisa FP Jr, et al. The implications of constraint in lumbar total disc replacement. *J Spinal Disord Tech* 2003;16:412–7.

# Spinal Cord Stimulation for Axial Low Back Pain

## A Prospective, Controlled Trial Comparing Dual With Single Percutaneous Electrodes

Richard B. North, MD,\* David H. Kidd, MA,\* John Olin, PA,\* Jeffrey M. Sieracki, MS,\*  
Farrokh Farrokhi, MD,\* Loredana Petrucci, MS,\* and Protagoras N. Cutchis, MD†

**Study Design.** A prospective, controlled, clinical trial comparing single and dual percutaneous electrodes in the treatment of axial low back pain from failed back surgery syndrome.

**Objectives.** To clarify technical requirements and test the hypothesis that placing two linear arrays in parallel, thereby doubling the number of contacts, improves outcome.

**Summary of Background Data.** Technical improvements have enhanced outcomes of spinal cord stimulation for chronic axial low back pain. Dual, parallel electrodes reportedly improve these outcomes.

**Methods.** Acting as their own controls, 20 patients who passed screening with single, 4-contact electrodes received permanent dual, 4-contact electrodes with 7- or 10-mm intercontact distances at the same vertebral level(s). We quantified and compared the technical and clinical results of the single and dual electrodes, adjusting stimulation parameters to specific psychophysical thresholds.

**Results.** Single electrodes provided significant ( $P < 0.01$ ) advantages in patient- and computer-calculated ratings of pain coverage by paresthesias and in the scaled amplitude necessary to cover the low back, compared with dual 7-mm electrodes. Slight advantages without statistical significance were observed for the single over the dual 10-mm electrodes. Amplitude requirements were significantly lower for the single electrode than for either dual electrode. At long-term follow-up, 53% of patients met the criteria for clinical success.

**Conclusions.** While we observed disadvantages for dual electrodes in treating axial low back pain, we achieved technical success with single or dual electrodes in most patients and maintained this success clinically with dual electrodes in 53%.

**Key words:** low back pain, failed back surgery syndrome, electrical stimulation, spinal cord stimulation. **Spine 2005; 30:1412–1418**

but many clinicians have reserved SCS to treat predominantly radicular rather than axial low back pain (LBP) because of the technical difficulty in achieving pain coverage by paresthesia in the lower back.<sup>1–14</sup>

The development of programmable multicontact devices and of patient-controlled computer stimulator adjustment as well as the refinement of techniques for achieving low back coverage have made it feasible to use SCS to treat predominantly LBP accompanied by secondary radicular pain.<sup>2,3,15–17</sup> To clarify the technical requirements for this application and test the hypothesis that increasing the number of contacts improves outcome, we conducted a prospective, controlled study that compared results using single versus dual electrodes.

### Materials and Methods

**Patient Selection.** We recruited all eligible patients referred to the lead author at the Johns Hopkins Hospital from 1994 to 1997 and obtained informed consent from all participants.

Inclusion required (1) diagnosis of FBSS with predominant LBP (see Figure 1)<sup>18,19</sup> and secondary radicular pain (to facilitate screening); (2) an abnormality (e.g., lumbar arachnoid fibrosis) revealed by a recent imaging study, a consistent neurologic deficit, and/or documented history of an operation for appropriate indications; and (3) failure of all conservative alternatives.

We applied standard SCS exclusion criteria (see Table 1).

**Methods.** We screened all patients using a 4-contact ( $1 \times 4$ ) percutaneous electrode with 9-mm intercontact separation and a removable, malleable stylet that facilitates localization of the midline (Medtronic 3487A, Medtronic, Inc., Minneapolis, MN) and our own patient-interactive computerized system for SCS adjustment.<sup>14,20,21</sup>

We positioned the  $1 \times 4$  in the radiographic midline under fluoroscopic guidance and delivered test stimulation with a single anode immediately caudad to a single cathode.<sup>22</sup> Bipolar stimulation at representative contact combinations along the length of the array established acceptable physiologic midline placement, which achieved an amplitude threshold for bilateral paresthesia no greater than 110% of the perceptual threshold (see Figure 2). If we failed to achieve the longitudinal target of low back coverage without discomfort at the lowest amplitude, we positioned the electrode using a cathode as the most cephalad contact so paresthesia would extend just below the umbilicus. Typical optimal electrode positions spanned T9–T10, and the screening lasted at least 3 days to establish the degree of pain relief.

Our patients quantified  $1 \times 4$  and  $2 \times 4$  stimulator performance on a computerized graphical interface by entering a pain drawing<sup>14,20,21</sup> and adjusting the amplitude to test different

Spinal cord stimulation (SCS) is an increasingly successful therapy for “failed back surgery syndrome” (FBSS),

From the \*Department of Neurosurgery, School of Medicine and †the Applied Physics Laboratory, Johns Hopkins University, Baltimore, Maryland.

Supported in part by Advanced Neuromodulation Systems, Inc. (formerly Neuromed, Inc., formerly Quest, Inc.), Allen, Texas.

Acknowledgment date: May 14, 2004. First revision date: July 21, 2004. Acceptance date: July 23, 2004.

The device(s)/drug(s) is/are FDA-approved or approved by corresponding national agency for this indication.

Corporate/Industry funds were received in support of this work. No benefits in any form have been or will be received from a commercial party related directly or indirectly to the subject of this manuscript.

Address correspondence and requests for reprints to Richard B. North, MD, Department of Neurosurgery, Johns Hopkins University School of Medicine, 600 N. Wolfe St., Meyer 8-181 Baltimore, MD 21287-7881; E-mail: RNorth@jhmi.edu

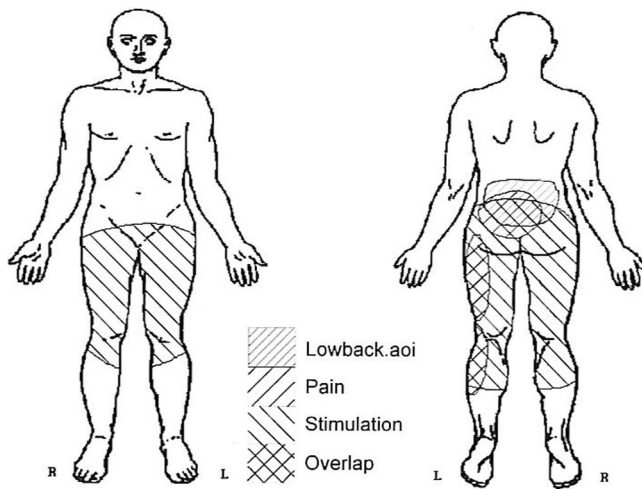


Figure 1. Drawing on a template modified from the 1949 Palmer original used in the McGill pain questionnaire illustrates our inclusion criterion of “axial low back pain” in the presence of radicular lower extremity pain. In this example, the technical goal of covering pain with stimulation paresthesia is achieved almost completely.

contact combinations with a 500 microsecond pulse width at 60 pulses per second.

We presented the 50 unique combinations of contact assignments possible with a  $1 \times 4$  in random order and blinded fashion, and, for each, the patients determined a “usage” amplitude (that which elicited a broad area of comfortable paresthesia), rated pain coverage or coverage by paresthesia on a visual analog scale, and drew the paresthesia to allow computerized calculation of coverage. The minimum criterion for SCS implantation was 50% pain relief validated by stable or reduced analgesic use and physical activity commensurate with physical condition.

In the first group of 10 patients, we implanted a dual, 4-contact ( $2 \times 4$ ) electrode [1994JF, Advanced Neuromodulation Systems (ANS, formerly Neuromed, formerly Quest), Plano, TX] with a 10-mm intercontact separation, and for the second group of 10 patients, we implanted a  $2 \times 4$  (2093-2193S, ANS) with the then newly available 7-mm intercontact separation. The  $1 \times 4$  and  $2 \times 4$  contacts had identical diameters and lengths. We positioned the  $2 \times 4$  at the same radiographic level as the  $1 \times 4$  (Figure 3) and sought to achieve bilateral paresthesia at an amplitude under 130% of perception without eliciting radicular thoracic paresthesia or motor response. The columns straddled the midline and were offset longitudinally by one-half the intercontact distance (5 mm or 3.5 mm)<sup>2,16,17</sup> and

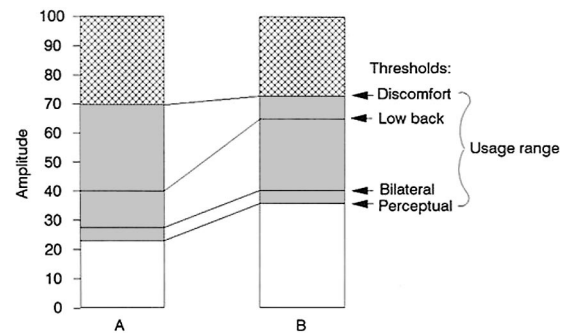


Figure 2. The amplitude range from perception to discomfort defines the clinically useful range of stimulation. We can scale the amplitudes at which a patient achieves bilateral effects and low back coverage with paresthesia to this range. Configurations **A** and **B** are comparably symmetric because they have similar ratios of bilateral to perceptual thresholds. Of the two, **A** achieves low back coverage closest to the perceptual threshold, and **B** achieves it nearest the discomfort threshold.

were separated horizontally by no more than one contact width.

Each  $2 \times 4$  array permitted 6050 unique combinations; this precluded exhaustive testing. Thus, for comparability with the  $1 \times 4$  tests, the patients tested the 45 randomized, blinded  $2 \times 4$  combinations that had at least one cathode 0.5 to 1 row above an anode in the same or an adjacent column (Figure 4). This also recognized that narrowly separated bipoles are more likely to recruit midline than lateral dorsal column fibers or dorsal roots (Figures 5 and 6).<sup>22,23</sup> For each combination, the patient noted the perceptual, usage, and discomfort amplitudes and rated the coverage of pain by paresthesia at the usage amplitude.

We retested the 10 combinations with the best pain coverage for each  $1 \times 4$  and  $2 \times 4$  configuration, adjusting each combination to discern the thresholds for perception, bilateral paresthesia, discomfort, usage, and paresthesia covering the low back. We then scaled the low back amplitude from perception (0) to discomfort (100).<sup>24</sup> Results  $>100$  indicated a lack of low back coverage below discomfort threshold.

For the patients receiving 7-mm  $2 \times 4$  electrodes, we also tested “guarded cathode” configurations<sup>2,23</sup> by adding anodes immediately cephalad to the row of cathodes for the 10 combinations with the highest coverage ratings. We randomized the order of presentation of these configurations and presented them to the patient in double-blind fashion, intermingled with “unguarded” configurations at all five thresholds (Table 2 summarizes the protocol).

**Outcome Assessment.** Our technical outcomes were patient-rated and computer-calculated measures of coverage and quantitative amplitude comparisons.

A disinterested third party, never involved in patient care, collected six-month, postimplantation follow-up outcome data via telephone and, in some cases, through interviews using a questionnaire that we mailed in advance and that patients returned after the interview (for validation and for analysis of visual analog scales). The questionnaire incorporated standard, validated measures of pain,<sup>25,26</sup> pain relief,<sup>25,26</sup> pain intensity quantified as a function of time,<sup>19</sup> quality of life,<sup>19</sup> activities of daily living,<sup>19,27</sup> functional capacity,<sup>19</sup> work status,<sup>19</sup> medication use,<sup>19</sup> pain experience,<sup>19</sup> and patient satisfaction with

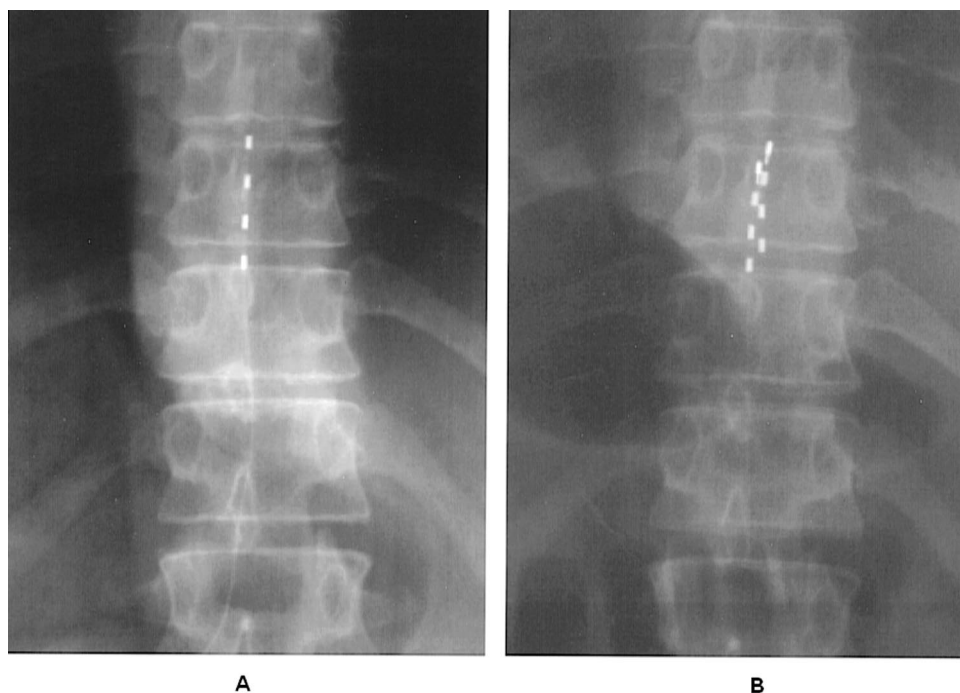
**Table 1. Spinal Cord Stimulation Exclusion Criteria**

1. A physiologic contraindication (e.g., a chief complaint of mechanical low back pain)
2. Abnormal pain behavior
3. A coexisting chronic pain problem or neurologic disease
4. A coexisting condition that would increase procedural risk (e.g., sepsis, coagulopathy)
5. A history of laminectomy or posterior fusion at the thoracolumbar junction, where percutaneous electrodes are routinely inserted
6. Unresolved psychiatric illness
7. Unresolved issues of secondary gain
8. Inappropriate medication use

Note. Used with permission.



Figure 3. The  $2 \times 4$  electrode is positioned at the same radiographic level as the  $1 \times 4$  (A). Each column of contacts in the  $2 \times 4$  (B) is positioned to one side of the midline, with no more than one contact width separating the two columns and an offset of one-half the intercontact distance.



treatment ("Considering the hospitalizations, discomfort, expense, would you go through it all again for the same result?").<sup>3,28-31</sup> All study patients were also interviewed when the last study patient reached the first postimplantation anniversary.

We defined long-term success as at least 50% continued pain relief combined with patient satisfaction with treatment.

**Statistical Methods.** A sample of 10 patients providing the 10 best combinations for analysis in each group was chosen to detect a difference of a 0.4 standard deviation with 80% statis-

tical power (1-beta) at a 0.05 level of significance. All *P* values are two-sided.

We averaged the best results for each patient for each configuration and used simple, descriptive, standard, paired and unpaired *t* tests to compare the 10 best  $1 \times 4$  and  $2 \times 4$  results for pain rating, coverage, amplitude usage thresholds, and scaled low back amplitude.

Because of the high level of inpatient correlation and the small sample size, we conducted an analysis of variance using patient and device identifiers as separate variables.

## ■ Results

All 10 patients in the first group proceeded to uneventful implantation of 10-mm  $2 \times 4$ s. In the second group, all 13 screened patients reported paresthesia coverage of pain, but 2 did not obtain adequate relief. Thus, 11 pro-

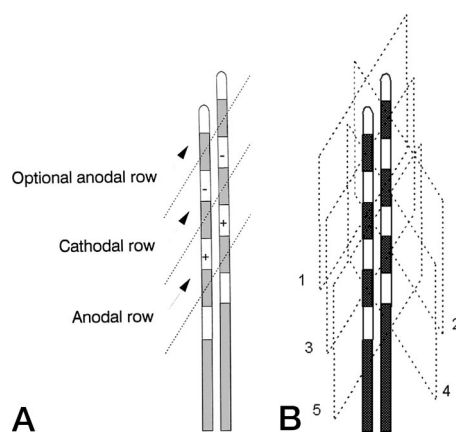


Figure 4. A, for the  $2 \times 4$ s, we assigned adjacent contacts on each electrode as cathodes (cephalad) or anodes (caudad), with at least one of each in each "row." In secondary tests, we added anodes to the row of contacts immediately cephalad (when such a row existed) and tested all combinations exhaustively for five quartets of contacts in adjacent diagonal rows. (No cephalad anodes could be added to quartet 1.) B, the offset design of the two columns creates oblique, alternating "rows" for five quartets of contacts, as shown here. Each quartet has nine potential contact combinations that shift the symmetry of stimulation.

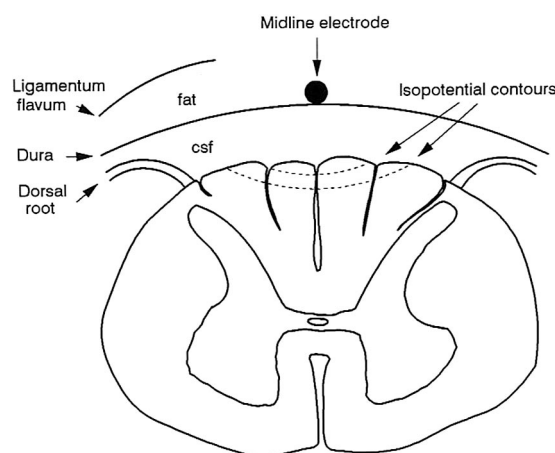


Figure 5. This cross-section of the spinal canal illustrates midline electrode placement and the concentric isopotential contours produced by spinal cord stimulation.

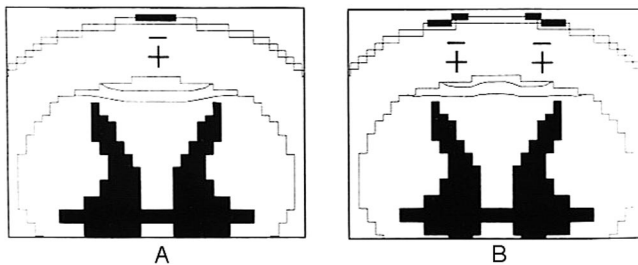


Figure 6. A finite element computer model indicates that electrical fields produced by dual electrodes do not sum constructively in the midline to increase the depth at which stimulation is effective. Given a cerebrospinal fluid thickness of 2.6 mm or less and electrode(s) placed at T8–T9, the single midline cathode (A) activates a deeper area with a lower amplitude laterally than do the side-by-side cathodes (B). Reproduced with permission from: Holsheimer J, Wesseling WA. Effect of anode-cathode configuration on paresthesia coverage in spinal cord stimulation. *Neurosurgery* 1997;41:654–60.

ceeded to uneventful implantation of 7-mm  $2 \times 4$ s, but one could not travel to the hospital to participate in the study.

Our 12 male and 8 female study participants, aged  $47.2 \pm 11.3$  years (range 31–72), had an average history of  $3.8 \pm 1.5$  (range 1–7) operations.

#### Technical Outcome

The  $1 \times 4$ s and 10-mm  $2 \times 4$ s did not differ significantly in areas of coverage or extraneous stimulation. All non-significant differences, however, favored the  $1 \times 4$ s (see Table 3). The 10-mm  $2 \times 4$  did not improve symmetry of stimulation coverage and required an average 53% absolute amplitude setting to achieve low back coverage, which was 50% more power than the average 43% setting needed for the  $1 \times 4$ .

We found a highly significant advantage for the  $1 \times 4$  versus the 7-mm  $2 \times 4$  in patient-rated ( $P = 0.003$ ) and computer-calculated area of coverage ( $P = 0.007$ ). As shown in Table 4, neither the computer calculated area of extraneous stimulation nor the symmetry of coverage differed. The analysis of variance showed no significant difference in repeated measures for individual patients, devices, or device settings.

The ratio of discomfort to perceptual threshold (see Figure 2) was  $1.75 \pm 0.36$  (range 1.17–3) for the  $1 \times 4$  and  $2.37 \pm 1.59$  (range 1.2–8.33) for the  $2 \times 4$ .

Table 2. The Research Protocol

All patients:
$1 \times 4$ configuration
50 contact combinations at usage threshold
10 best combinations at all five thresholds
All patients who passed the trial:
$2 \times 4$ configuration
45 contact combinations at perception, bilateral paresthesia, and usage thresholds
10 best combinations at all five thresholds
Patients with (7 mm inter-contact distance electrodes):
10 best combinations (modified to guarded configurations) at all five thresholds

Table 3. Quantitative Comparison of  $1 \times 4$  and 10 mm  $2 \times 4$  Configurations for 10 Patients

	Coverage Rating	Coverage Area	Extraneous Area	Symmetry Score	Pulse Amplitude
$1 \times 4$	$63 \pm 26$	$53 \pm 26$	$55 \pm 23$	$82 \pm 24$	$21 \pm 9$
$2 \times 4$	$61 \pm 31$	$52 \pm 30$	$57 \pm 24$	$78 \pm 27$	$32 \pm 10$

Note. All figures are mean  $\pm$  SD; all variables are percentages (0–100 scale). The symmetry score represents  $100 - 2 \times (|\% \text{ of pixels on left side} - 50|)$ .

Finally, the guarded configurations for the best 10 combinations, compared with the original best 10, showed no advantage. In fact, the calculated coverage at the usage threshold was slightly better ( $62 \pm 28$ ,  $n = 54$  vs.  $55 \pm 29$ ,  $n = 64$ ) at a slightly lower amplitude ( $29 \pm 16$  vs.  $36 \pm 29$ ) for the unguarded configurations.

#### Clinical Outcome

At 2.3-year average follow-up (range 1.1–4.2), the 19 available patients (95%) reported 46% average relief of LBP. Of 17 patients who provided data on change in pain relief from implantation to follow-up, 10 (59%) reported no change (2 of these had SCS explanted; 2 had stopped use), 5 (29%) less relief, and 2 (12%) an increase in pain relief (both had SCS explanted).

Of 13 patients who continued stimulation, 10 (53%) met both criteria for success. Coverage of radicular pain was consistently better than coverage of LBP, and coverage of both declined only slightly over time (see Table 5) as did relief of LBP (see Table 6), which was less than relief of radicular pain.

Most patients reported no pain impairment in performing activities of daily living, but stimulator use decreased strength or coordination in 9 (53%), sense of touch in 8 (47%), and bladder or bowel control in 2 (12%). Of the 17 patients supplying data about medication, 7 (41%, including 6 successes) decreased and 9 (53%) increased their medication use, and one stayed the same (see Figure 7).

Of 17 (85%) patients employed before their illness, 8 (40%) were working immediately before SCS implantation and 7 (35%) at long-term follow-up.

#### Complications

Electrode migration in one patient and receiver failure in another required surgical revision. Clinical failure led to the removal of implants in 4 patients (20%) after an average of 0.9 years (range 0.5 to 1.4).

Table 4. Quantitative Comparison of  $1 \times 4$  and 7 mm  $2 \times 4$  Configurations for 10 Patients

	Coverage Rating	Coverage Area	Extraneous Area	Symmetry Score	Pulse Amplitude
$1 \times 4$	$67 \pm 27$	$65 \pm 29$	$64 \pm 27$	$55 \pm 21$	$20 \pm 30$
$2 \times 4$	$58 \pm 26$	$57 \pm 28$	$65 \pm 27$	$57 \pm 20$	$44 \pm 31$
<i>P</i>	0.0032	0.0071	ns	ns	0.0000

Note. All figures are mean  $\pm$  SD; all variables are percentages (0–100 scale). The symmetry score represents  $100 - 2 \times (|\% \text{ of pixels on left side} - 50|)$ .

**Table 5. Percentage of Patients Reporting Stimulation Coverage of Low Back Pain**

	6 Months n = 15	2 Years n = 12
Axial	73	67
Radicular	100	92

## ■ Discussion

We have reported previously that in FBSS patients, SCS led to comparable relief of primary radicular pain and secondary LBP.<sup>3</sup> Our present study patients reported less coverage and less relief of primary LBP than of secondary radicular pain, perhaps partly because LBP has a nociceptive component (the value of SCS is better established for neuropathic pain).<sup>32,33</sup> Nevertheless, we achieved clinical success comparable with that in our previous studies,<sup>4</sup> and our results support expanding SCS's indication to patients with predominant LBP.

Two retrospective/follow-up studies underscore our clinical findings. One involving 17 FBSS patients with low back and leg pain found that treatment with  $2 \times 4$  electrodes significantly reduced global, back, and leg pain as well as analgesic use.<sup>34</sup> The other examined outcome in 41 patients and found that 60% considered themselves improved and 75% would undergo the procedure for the same result.<sup>35</sup> In a prospective study, investigators used paddle electrodes to treat 44 patients with LBP predominant or comparable with leg pain and found that leg pain relief exceeded back pain relief at 6 and 12 months, with the gap widening at 12 months.<sup>36</sup> None of these studies compared dual electrodes with other electrode configurations.

We hypothesized that multicontact electrodes, which allow noninvasive adjustment,<sup>4,13</sup> would be more likely than single-contact arrays to compensate for asymmetry or variability in pain patterns; however, we observed no technical advantage in using the  $2 \times 4$ s. In fact, we found a statistically significant disadvantage in achieving coverage as well as a significantly increased power requirement, which of course reduces battery life.<sup>37</sup>

This lack of advantage is remarkable because in our previous similarly-designed study of FBSS patients with primarily radicular pain, we randomly implanted either the same type of  $1 \times 4$  percutaneous electrode used in the screening trial or a  $1 \times 4$  insulated array with a laminotomy paddle configuration, and the implant outcome in each group surpassed the trial results.<sup>38,39</sup> Thus, the present results provide our first indication that any permanently implanted configuration can be inferior to a temporary percutaneous  $1 \times 4$ .

The present study offers no support for the claim that symmetrically placed, parallel electrodes with each contact positioned, as designed, off midline by at least half its width create an advantageous midline "summation effect." Indeed, one would not expect this theoretically: assuming a simple inverse square relationship between

voltage and distance from the stimulating cathode, the electrical fields between two cathodes do not sum constructively in the midline. Computerized finite element modeling<sup>40</sup> confirms this (see Figure 6).

Another potential advantage of using parallel electrodes to treat LBP might be the possibility of controlling symmetry of paresthesia. Although our study shows we can adjust symmetry with dual electrodes, we found no technical advantage over a single electrode placed in the physiologic midline. The manufacturer's recommendation of separating each contact by half the intercontact distance<sup>2,16,17</sup> advantageously shortens the dipoles but results in broad, irregular, oblique, "virtual" contacts, which computer models predict (and our results corroborate) are inferior to a zero contact offset.<sup>36</sup>

Our results suggest that 1.2 mm-wide contacts do not provide sufficient horizontal resolution to compensate for off-midline placement. Narrower electrodes might yield different results. Alternatively, it is possible to bracket a single midline electrode with two electrodes;<sup>2</sup> such a triple electrode system, however, with electrodes inserted and anchored independently, would be cumbersome and might increase the likelihood of migration. Investigators have described a "transverse tripolar," two-dimensional, implanted array but found it of no value in treating LBP.<sup>41</sup>

Despite our objective of treating LBP, all of our patients had lower extremity pain, the symmetry of which varied. A potential advantage of a  $2 \times N$  electrode is the ability to direct paresthesia to the right or left. The  $1 \times 4$  does not offer this advantage, yet allowed us to gain better overlap.

Devices that deliver interleaved or simultaneous pulse trains using different contact combinations (e.g., to the left and right or proximally and distally), however, might create a time-related rather than space-related "summation effect"<sup>14,17,22,42,43</sup> that could improve treatment of LBP.

The increased amplitude requirement we observed for dual electrodes partly results from the smaller intercontact distance of staggered contacts. This technical problem can be managed (e.g., by using an external radiofrequency power source), but we found no clinical advantage and some disadvantages for dual electrodes with the amplitude scaled to the clinically relevant range. Our observations for dual electrodes, however, are limited to configurations with offset contacts.

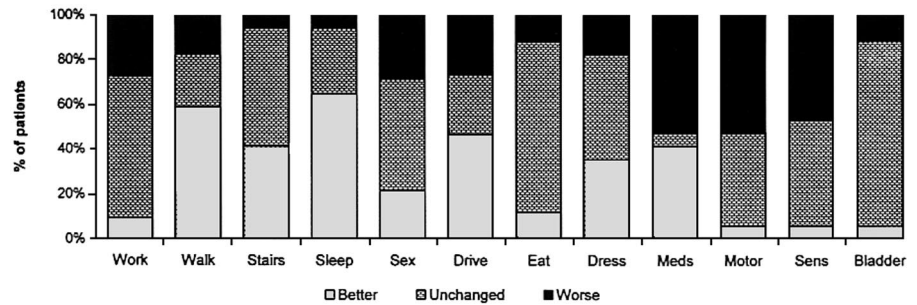
A longer  $2 \times 8$  system might increase the likelihood of spanning the optimal contact position. In uncontrolled studies,  $2 \times 8$  systems with 7-mm spacing have been

**Table 6. Percentage of Patients Reporting Relief of Low Back Pain**

	6 Months n = 16	2 Years n = 13
Axial	50	46
Radicular	88	92



Figure 7. Histogram of the percentage of patients reporting changes in (1) pain impairment of activities of daily living, (2) medication use, and (3) neurologic symptoms. Most patients reported no change in everyday activities or in medication use.



useful in the treatment of LBP.<sup>2,16,17</sup> A  $2 \times 8$  system, however, is simply two  $2 \times 4$  systems placed end-to-end,<sup>24</sup> and the available systems have identical intercontact spacing. Assuming that an experienced implanter can identify the optimal longitudinal target in a naive patient and that our  $1 \times 4$  and  $2 \times 4$  spanned the optimal longitudinal position, we might expect the same results in a study of  $1 \times 4$ s or  $1 \times 8$ s versus  $2 \times 8$ s.

### Study Limitations

In this, our first study of the use of SCS for primarily LBP, each participant acted as his/her own control, and the temporary  $1 \times 4$  was compared with a permanent  $2 \times 4$ . A prospective, randomized, longitudinal comparison of  $1 \times 4$  and  $2 \times 4$  permanent electrodes is needed to determine definitively which offers the best outcome.

Although we did not collect computerized measurements to quantify long-term technical results, our patients reported maintenance of low back coverage and pain relief. No long-term measures are available, of course, for our  $1 \times 4$  screening electrodes.

Despite the facts that the 91% SCS trial success rate in this study was our highest reported to date and that those who failed to achieve sufficient pain relief nevertheless achieved paresthesia coverage, 20% (4 of 20) had their devices explanted, and another 2 stopped use because of inadequate pain relief. The 20% explantation rate, unusually high in our experience but not significant given the small sample size, might reflect an initial confusion between radicular stimulation and low back coverage on the part of our SCS-naïve patients. We have not encountered this problem before, however, even though only 1 to 3 days routinely pass between naïve  $1 \times 4$  screening and  $2 \times 4$  technical tests.

We might have underutilized the  $2 \times 4$  system by limiting the testing of available contact combinations to the reportedly advantageous dipoles of minimum length with cathodes immediately above anodes.<sup>24</sup> We did, however, test the  $2 \times 4$  configurations with the guarded cathode geometries that, for  $1 \times 4$ s in patients with primarily radicular pain, have a statistically significant clinical<sup>13</sup> and psychophysical<sup>22</sup> association with improved coverage. Although we tested fewer total  $1 \times 4$  combinations, the  $1 \times 4$ s nevertheless provided the best technical outcome.

Furthermore, the preference for  $1 \times 4$  configurations did not include the disproportionate number of guarded combinations we observed among patients in prior stud-

ies.<sup>13,22</sup> Perhaps guarded configurations, which explained a relatively small amount of variance in our prior studies, offer no advantage to patients with LBP.

Although our sample was adequate to detect clinically important technical differences, a larger sample might have revealed additional differences.

Finally, the precise midline placement of  $1 \times 4$  electrodes achieved in a consecutive series of patients by or under the supervision of a surgeon who has placed more than 2000 percutaneous electrodes does not represent general clinical practice. With inexperienced implanters, it might be easier to demonstrate the potential of dual electrodes to compensate for imprecise placement.

### Conclusions

We observed no technical advantage for implanted  $2 \times 4$  electrodes compared with percutaneous  $1 \times 4$  electrodes in the same patients at the same level for the treatment of LBP. Compared with the  $1 \times 4$  electrode, the  $2 \times 4$ s with 10-mm intercontact spacing were slightly but not significantly inferior (except for their significantly greater power requirement). The  $2 \times 4$ s with 7-mm intercontact spacing had significantly inferior patient and computerized ratings of coverage as well as increased power consumption.

Long-term clinical follow-up, however, revealed that SCS with dual electrodes was successful in a majority of patients with a chief complaint of LBP.

Future studies should compare the results obtained with other electrode configurations and pulse parameters, including interleaved pulse trains using multiple contact combinations.

### Key Points

- Technical advantages have made it possible to treat axial low back pain with spinal cord stimulation.
- The use of dual, parallel electrodes for spinal cord stimulation reportedly improves outcome.
- Quantifying and comparing technical and clinical results using single and dual electrodes revealed disadvantages for dual electrodes in treating axial low back pain.
- We achieved technical success treating low back pain with single and dual electrodes in most patients and maintained this success clinically in 53%.

## References

1. Law JD. Targeting a spinal stimulator to treat the "failed back surgery syndrome." *Appl Neurophys* 1987;50:437-8.
2. Law JD, Kirkpatrick AF. Pain management update: spinal cord stimulation. *Am J Pain Management* 1991;2:34-42.
3. North RB, Ewend MG, Lawton MT, et al. Failed back surgery syndrome: five-year follow-up after spinal cord stimulator implantation. *Neurosurgery* 1991;28:692-9.
4. North RB, Kidd DH, Zahurak M, et al. Spinal cord stimulation for chronic, intractable pain: experience over two decades. *Neurosurgery* 1993;32:384-95.
5. Burchiel KJ, Anderson VC, Brown FD, et al. Prospective, multicenter study of spinal cord stimulation for relief of chronic back and extremity pain. *Spine* 1996;21:2786-94.
6. Ohnmeiss DD, Rashbaum RF, Bogdanffy GM. Prospective outcome evaluation of spinal cord stimulation in patients with intractable leg pain. *Spine* 1996;21:1344-51.
7. Nashold B, Somjen G, Friedman H. Paresthesias and EEG potentials evoked by stimulation of the dorsal funiculi in man. *Exp Neurol* 1972;36:273-87.
8. Sweet WH, Wepsic JG. Electrical stimulation for suppression of pain in man. In: Fields WS, editor. *Neural Organization and Its Relevance to Prosthetics*. New York, NY: Intercontinental Medical Book Corporation, 1973:218-40.
9. Burton CV. Session on spinal cord stimulation: safety and clinical efficacy. *Neurosurgery* 1977;1:164-5.
10. Urban BJ, Nashold B. Percutaneous epidural stimulation of the spinal cord for relief of pain: long term results. *J Neurosurg* 1978;48:323-8.
11. Krainick JU, Thoden U. Dorsal column stimulation. In: Wall PD, Melzack R, editors. *Textbook of Pain*. New York, NY: Churchill Livingstone, 1989: 701-5.
12. Kumar K, Nath R, Wyant GM. Treatment of chronic pain by epidural spinal cord stimulation: a 10-year experience. *J Neurosurg* 1991;75:402-7.
13. North RB, Ewend MG, Lawton MT, et al. Spinal cord stimulation for chronic, intractable pain: superiority of "multichannel" devices. *Pain* 1991; 44:119-30.
14. North RB, Fowler KR, Nigrin DA, et al. Patient-interactive, computer-controlled neurological stimulation system: clinical efficacy in spinal cord stimulation. *J Neurosurg* 1992;76:689-95.
15. North RB, Kidd DH, Lee MS, et al. Spinal cord stimulation versus reoperation for the failed back surgery syndrome: a prospective, randomized study design. *Stereotact Funct Neurosurg* 1994;62:267-72.
16. Rossi U. Technical advances in neuromodulation: state of the art hardware technology in neurostimulation. *Abstracts of the 3rd International Congress of the Neuromodulation Society*, 1996 March 6-10. Orlando, FL: San Francisco, CA: International Neuromodulation Society, 1996.
17. Alo KM, Yland MJ, Kramer DL, et al. Computer assisted and patient interactive programming of dual octrode spinal cord stimulation in the treatment of chronic pain. *Neuromodulation* 1998;1:30-45.
18. Palmer H. Pain charts: a description of a technique whereby functional pain may be diagnosed from organic pain. *N Z Med J* 1949;264:187-213.
19. Melzack R. The McGill Pain Questionnaire: major properties and scoring methods. *Pain* 1975;1:277-99.
20. North RB, Nigrin DA, Fowler KR, et al. Automated "pain drawing" analysis by computer-controlled, patient-interactive neurological stimulation system. *Pain* 1992;50:51-8.
21. North RB, Sieracki JM, Fowler KR, et al. Patient-interactive, microprocessor-controlled neurological stimulation system. *Neuromodulation* 1998;1: 185-93.
22. Law JD. Spinal stimulation: statistical superiority of monophasic stimulation of narrowly separated, longitudinal bipoles having rostral cathodes. *Appl Neurophys* 1983;46:129-37.
23. Holsheimer J, Wesselink WA. Optimum electrode geometry for spinal cord stimulation: the narrow bipole and tripole. *Med Biol Eng Comput* 1997;35: 493-7.
24. Law JD. A new method for targeting a spinal stimulator: quantitatively paired comparisons. *Appl Neurophys* 1987;50:436.
25. Downie WW, Leatham PA, Rhind VM, et al. Studies with pain rating scales. *Ann Rheum Dis* 1978;37:378-81.
26. Jensen MP, Karoly P, Braver S. The measurement of clinical pain intensity: a comparison of six methods. *Pain* 1986;27:117-26.
27. Sternbach RA, Wolf SR, Murphy RW, et al. Traits of pain patients: the low-back "loser." *Psychosomatics* 1973;14:226-9.
28. Law JD, Lehman RAW, Kirsch WM. Reoperation after lumbar intervertebral disc surgery. *J Neurosurg* 1978;48:259-63.
29. Finnegan WJ, Fenlin JM, Marvel JP, et al. Results of surgical intervention in the symptomatic multiply-operated back patient. *J Bone Joint Surg* 1979;61: 1077-82.
30. Waddell G, Kummel EG, Lotto WN, et al. Failed lumbar disc surgery and repeat surgery following industrial injuries. *J Bone Joint Surg* 1979;61: 201-7.
31. De la Porte C, Siegfried J. Lumbosacral spinal fibrosis (spinal arachnoiditis): its diagnosis and treatment by spinal cord stimulation. *Spine* 1983;8:593-603.
32. Sánchez-Ledesma MJ, Garcia-March G, Diaz-Cascajo P, et al. Spinal cord stimulation in deafferentation pain. *Stereotact Funct Neurosurg* 1989;53: 40-55.
33. Marchand S, Bushnell MC, Molina-Negro P, et al. The effects of dorsal column stimulation on measures of clinical and experimental pain in man. *Pain* 1991;45:249-57.
34. Van Buyten J-P, Zundert JV, Milbouw G. Treatment of failed back surgery syndrome patients with low back and leg pain: a pilot study of a new dual lead spinal cord stimulation system. *Neuromodulation* 1999;2:258-65.
35. Ohnmeiss DD, Rashbaum RF. Patient satisfaction with spinal cord stimulation for predominant complaints of chronic, intractable low back pain. *Spine* 2001;1:358-63.
36. Barolat G, Oakley JC, Law JD, et al. Epidural spinal cord stimulation with a multiple electrode paddle lead is effective in treating intractable low back pain. *Neuromodulation* 2001;4:59-66.
37. North RB, Brigham DD, Khalessi A, et al. Spinal cord stimulator adjustment to maximize implanted battery longevity: a randomized, controlled trial using a computerized, patient-interactive programmer. *Neuromodulation* 2004;7:13-25.
38. North RB, Kidd DH, Olin JC, et al. Spinal cord stimulation electrode design: prospective, randomized, controlled trial comparing percutaneous and laminectomy electrodes—Part I: technical outcomes. *Neurosurgery* 2002;51: 381-9.
39. North RB, Lanning A, Hessels R, et al. Spinal cord stimulation with percutaneous and plate electrodes: side effects and quantitative comparisons. *Neurosurgical Focus* 1997;2:1-5.
40. Holsheimer J, Wesselink WA. Effect of anode-cathode configuration on paresthesia coverage in spinal cord stimulation. *Neurosurgery* 1997;41:654-60.
41. Struijk JJ, Holsheimer J. Transverse tripolar spinal cord stimulation: theoretical performance of a dual channel system. *Med Biol Eng Comp* 1996;34: 273-9.
42. Holsheimer J, Nuttin B, King GW, et al. Clinical evaluation of paresthesia steering with a new system for spinal cord stimulation. *Neurosurgery* 1998; 42:541-7.
43. North RB, Fowler KF. Computer-controlled, patient-interactive, multichannel, implanted neurological stimulators. *Appl Neurophys* 1987;50:39-41.

# A New Operative Classification of Idiopathic Scoliosis: A Peking Union Medical College Method

Guixing Qiu, MD,\* Jianguo Zhang, MD,\* Yipeng Wang, MD,\* Hongguang Xu, MD,\*  
Jia Zhang, MD,\* Xisheng Weng, MD,\* Jin Lin, MD,\* Yu Zhao, MD,\* Jianxiong Shen, MD,\*  
Xinyu Yang, MB,\* Keith DK Luk, FRCS, FHKAM(Orth),† Duosai Lu, MD, PhD,† and  
William W. Lu, PhD†

**Study Design.** A retrospective radiographic study on the type of surgically treated idiopathic scoliosis, with a prospective study on the reliability of the type-related fusion guide.

**Objectives.** To identify and classify surgically treated idiopathic scoliosis, and define its related fusion levels by a new classification system.

**Summary of Background Data.** Some classification methods for idiopathic scoliosis have been suggested. However, poor intraobserver reproducibility and interobserver reliability were experienced in these studies, and were not appropriate for guiding surgical planning.

**Methods.** A total of 427 surgically treated idiopathic scoliosis cases were reviewed. Preoperative and postoperative standing anteroposterior, lateral, and preoperative supine side-bending radiograph were analyzed using the Scoliosis Research Society definition of scoliosis and curve apex. The resulting classification was tested for intraobserver reliability and interobserver reliability, and by 6 surgeons. Apical frequencies were determined for each type, and prospective surgical testing of the new type and its related fusion guide was performed.

**Results.** Three major types and 13 subtypes were identified, of which the Peking Union Medical College type I accounted for 56.62%, type II 42.16%, and type III 1.22%. The interobserver reliability testing was 85% (kappa coefficient 0.832), while intraobserver reproducibility was 91% (kappa coefficient 0.898). Each type had its corresponding fusion levels. A prospective study of 152 cases was performed according to the classification. All of these cases were followed over 18 months, and no postoperative decompensation was noted.

**Conclusion.** The Peking Union Medical College classification of idiopathic scoliosis is one system to combine each type with its corresponding fusion level, and it had much higher interobserver reliability and intraobserver reproducibility than the King system. Further prospective

studies would help to clarify and expand this system.

**Key words:** idiopathic scoliosis, classification, Peking Union Medical College. **Spine** 2005;30:1419–1426

Idiopathic scoliosis is a 3-dimensional (3-D) spinal deformity with unknown etiology. Surgical intervention is generally required if spinal deformities are severe or progressive. For the single curve such as thoracic, thoracolumbar, or lumbar curve, there are fewer differences in the selection of fusion level among different spinal surgeons except for the surgical approaches. However, the choice of fusion levels in some type of curves, such as double curves, remains a difficult and controversy issue. Inadequate fusion in these curves may result in postoperative curvature deterioration, trunk decompensation, or even produce new deformity.<sup>1–4</sup> The selection of approach and fusion level should be based on the inherent characteristics of the different curve types. A good classification system should include different types of the curves and should be a guide for surgical planning. Many investigators<sup>1,5–7</sup> have tried to develop an ideal classification system for idiopathic scoliosis. Although it is not an ideal system, the King classification is still the most widely referred system in the literature for surgical planning and comparing the results of different treatment. To establish a comprehensive and reliable system to help surgeons choose an appropriate approach and fusion level, 427 cases of surgically treated idiopathic scoliosis in our department were analyzed, and a new classification system called the Peking Union Medical College (PUMC) system was proposed. This system's reliability and reproducibility were tested.

## ■ Materials and Methods

**Clinical Materials.** From February 1983 to January 2001, 1245 patients with scoliosis were admitted and treated in our department. These patients were examined, and a computerized database was created. Of these patients, 427 (34.3%) were idiopathic and had undergone surgical treatment. The sex ratio was 1:2 (male-female), and the average age of the patients was 16.7 years. The higher ratio of male patients in this group may be a result of the distribution of the patients because most in our department were from the underdeveloped countryside, where more care and economic support from the parents came to the boys. Each case was analyzed by a measurement taken from the preoperative supine side-bending radiograph, anteroposterior and lateral standing radiographs taken before and

From the \*Department of Orthopedic Surgery, Peking Union Medical College (PUMC) Hospital, Beijing, China, and †Department of Orthopaedic Surgery, The University of Hong Kong, Hong Kong, China. Acknowledgment date: March 3, 2003. First revision date: May 5, 2003. Second revision date: April 16, 2004. Third revision date: June 11, 2004. Fourth revision date: August 1, 2004. Acceptance date: August 2, 2004.

Supported by the major project of Ministry of Health of the People's Republic of China.

The manuscript submitted does not contain information about medical device(s)/drug(s).

Foundation funds were received in support of this work. No benefits in any form have been or will be received from a commercial party related directly or indirectly to the subject of this manuscript.

Address correspondence and reprint requests to Guixing Qiu, MD, Department of Orthopaedic Surgery, Peking Union Medical College Hospital, No 1 Shuaifuyuan hutong, Beijing 100730, China; E-mail: qiugx@medmail.com.cn



after surgery. The second, third, and fourth authors (J.Z., Y.W., and H.X., respectively) of this article retrospectively measured the radiographs of 427 patients.

**Methods.** In our practice, we believe that the most confused issues are how to define a curvature and to what degree one curve is defined larger than the other. These issues play a key role in the selection of fusion in double and triple curves. To avoid such confusion and improve the reliability and reproducibility of this classification, we used the Scoliosis Research Society (SRS) definition of a curvature (*i.e.*, spine deviates from the midline with a Cobb angle more than 10°, and the furthest horizontal vertebra or disc from the midline on the standing anteroposterior view is defined as the apex).<sup>8</sup> Thus, curves were counted and classified into single, double, and triple curves according to the apex number. The frequency of each curvature type was also calculated.

The general principle of this classification is to preserve more mobile segment as possible, which is reasonable and essential to decrease the degenerative changes of the unfused segments. For this reason, we try to define the criteria for selective thoracic or selective lumbar fusion in double curves. Because a larger curve has poorer compensating ability, it is usually fused to let the smaller curve compensate for the spinal balance. Considering that the interobserver's variability ranges from 4.9° to 11.8°,<sup>9,10</sup> we define a larger curve and a minor curve as a 10° difference, and 2 equivalent curves if the difference is less than 10° to avoid a dispute. Therefore, in this system, selective fusion is fusing a larger curve or fusing the less flexible curve if the 2 curves are equal.

For each case, the following data were measured and recorded:

1. The Cobb angle of each curve.
2. Flexibility of the curvature: flexibility (%) =  $\frac{\text{Cobb angle on standing} - \text{Cobb angle on convex bending}}{\text{Cobb angle on standing}} \times 100\%$
3. Rotation of the apical vertebra: Rotation of vertebral body was recorded from 1° to 4° using the Nash-Moe method.<sup>11</sup>
4. The stable vertebra: Stable vertebra was defined as the first distal vertebra divided equally by the central sacrum vertical line defined by King et al.<sup>1</sup>
5. Thoracolumbar kyphosis: The Cobb angle between T12 and L1 is more than zero on the sagittal plane.

Based on the radiographic analysis, all the curves were divided into 3 major types according to the total number of the apex, and each major type is then further divided into its respective subtypes according to the location of the apex. Each subtype has its own morphologic pattern and appropriate vertebral fusion levels.

**Reliability and Reproducibility.** Roentgenographs from 29 cases were selected and numbered randomly by the first author (G.Q.), and classified independently by 6 other senior surgeons from the same department. One case was covered in type Ia, Ib, Ic, IIa, and III, considering the lower frequent of type IIa, III and the simplicity of Ia, Ib, and Ic. At least 3 cases of each subtype II, except for IIa, were covered to test the reliability and reproducibility of the relatively complicated subtypes of type II. All the marks on the films were cleaned before and after each surgeon's review, and the results were collected individually for statistical analysis. The interobserver reliability was analyzed

by comparing the results between surgeons. Two weeks later, the same surgeons were asked to repeat the classification using the same films but in a different order, and the results were used to analyze the intraobserver reproducibility. The kappa coefficient, which is between -1 and +1, calculated by coherence tests was used for the assessment. A coefficient more than zero indicates significance, and the higher the kappa coefficient, the better the coherence. A kappa coefficient  $\geq 0.75$  indicates an excellent level of coherence.

**Clinical Validation.** A prospective study was conducted on 152 patients who underwent surgical correction under the guidance of the PUMC classification and who were followed for more than 18 months until June 2002. Curve types and fusion levels were classified by the first 3 senior authors (G.Q., J.Z., and Y.W., respectively).

## ■ Results

### *Characteristics of the PUMC Classification System*

In this new PUMC classification system, spinal curvatures were divided into 3 main categories according to the number of apexes to be easily remembered: type I for 1, type II for 2, and type III for 3 apexes. There are a number of subtypes for each curve type, with a total of 13 subtypes (Figure 1). The frequency of each subtype from the 427 cases of idiopathic scoliosis is shown in Table 1.

### *Surgical Approach and Range of Fusion of Each Type*

Indication for surgical treatment includes curvature more than 40° and progression of 5° or more per 6 months except for conservative treatment. The third generation instrumentation is recommended in this system.

### *PUMC Type I: Single Curve*

**Subtype Ia.** Thoracic curve, apex between T2 and T11–T12 disc. In this group, 175 patients had this type (40.99%, Table 1). Of these patients, 107 had undergone only posterior fusion, 57 anterior release and posterior fusion, and 11 anterior arthrodesis and posterior fusion. When the curve flexibility is less than 50% or the curve is more than 50° on the bending film, anterior release is performed to obtain a better posterior correction. Anterior epiphysiodesis is indicated when the patient is far from skeletal maturity in this group. All the thoracic curves were fused to the stable vertebra (the most proximal vertebra which is bisected by the central sacrum vertical line is defined as the stable vertebra) and had a good coronal balance. Therefore, thoracic fusion to the stable vertebra is suggested for type Ia.

**Subtype Ib.** Thoracolumbar curve, apex at T12, T12/L1 disc, and L1. This type accounts for 7.73% (Table 1). Of the 33 patients, 22 had undergone only anterior fusion with instrumentation, 7 anterior release and posterior fusion with instrumentation, and only 4 underwent posterior fusion. Of the 22 patients with anterior fusion, a mean of 2.1 segments were saved compared with a standard posterior procedure. Seven of these 22 patients had fusion from end-to-end, and 15 had short segmental fu-

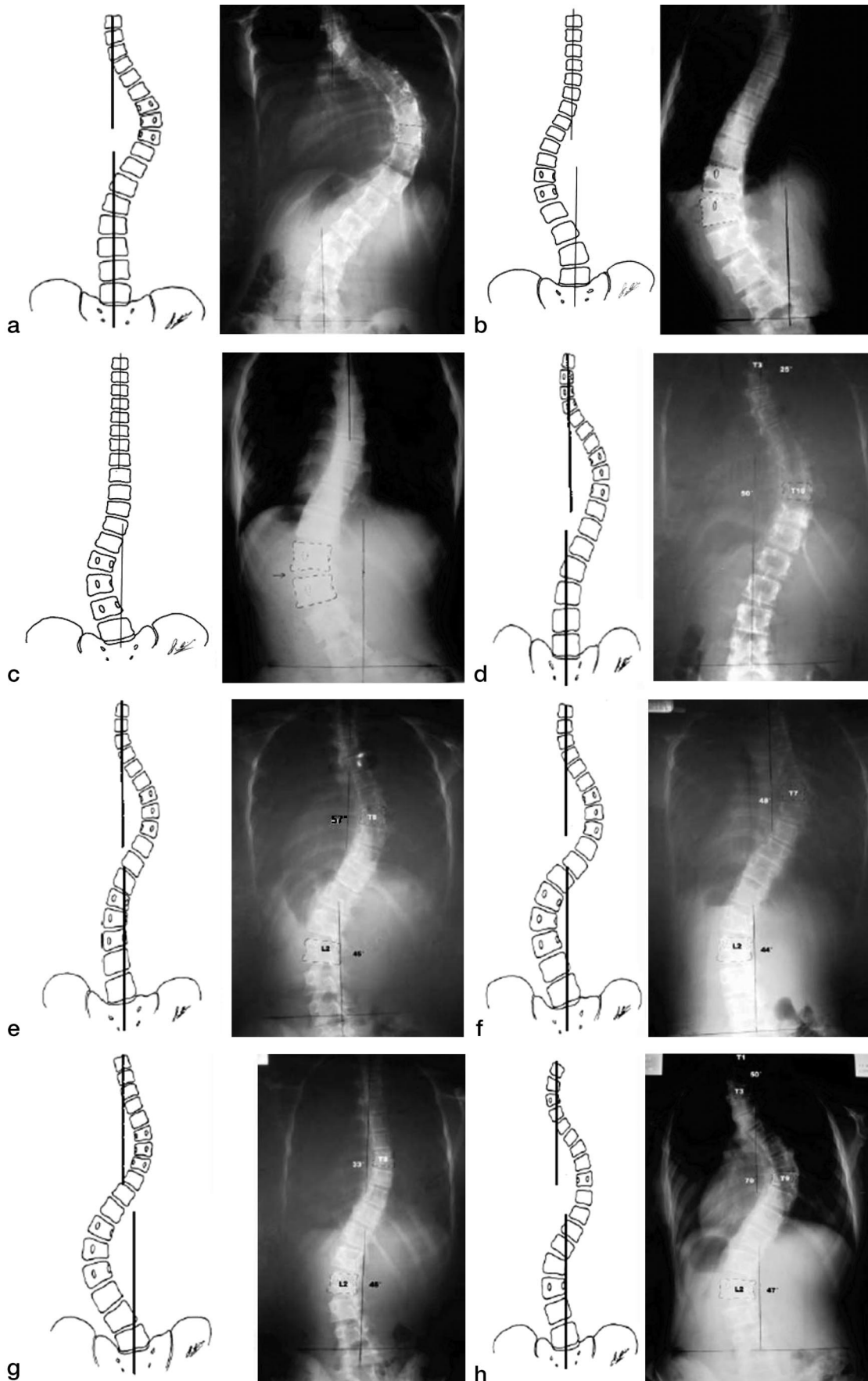


Figure 1. **a–h**, Schematic drawing and related radiograph films of PUMC classification. **a**, PUMC Ia. **b**, PUMC Ib. **c**, PUMC Ic. **d**, PUMC IIa. **e**, PUMC IIb. **f**, PUMC IIc. **g**, PUMC IId. **h**, PUMC III. The dotted line is the plumb line passing the spinous process of C7, and the black line is the center sacrum vertical line.

**Table 1. The Percentage of Each Type of PUMC Classification**

PUMC Type	Case No.	Percentage
I		
Subtype Ia	175	40.99
Subtype Ib	33	7.73
Subtype Ic	33	7.73
II		
Subtype IIa	13	3.05
Subtype IIb1	57	13.35
Subtype IIb2	51	11.94
Subtype IIc1	3	0.70
Subtype IIc2	3	0.70
Subtype IIc3	15	3.51
Subtype IId1	28	6.56
Subtype IId2	10	2.34
III		
Subtype IIIa	3	0.70
Subtype IIIb	3	0.70

sion according to the Hall principle,<sup>12</sup> and both coronal and sagittal balance were maintained, except that 2 patients with short segmental fusion had increased disc angle distal to the inferior fusion vertebra. A mean 13° of thoracolumbar kyphosis in 5 patients was corrected to 3.5°. For type Ib, we recommend anterior fusion with instrumentation except for cases with a curve more than 60° and flexibility less than 50%, or a curve more than 50° on the bending film. According to Hall *et al*,<sup>12</sup> the principle of selection of short segmental fusion levels is: (1) If the apex is a vertebra on standing anteroposterior (AP) film, instrument one vertebral body above and below; if a disc, instrument 2 vertebral bodies above and below. (2) On convex bending film, the first disc space above and below the apex that opens up can be left unfused; on concave bending film, vertebral bodies below the apex should be parallel to the sacrum. If there is a discrepancy among the levels indicated in the aforementioned 2 methods, the longest segment of instrumentation always should be selected.

**Subtype Ic.** Lumbar curve, apex between L1–L2 and L4–L5 intervertebral disc. A total of 33 patients were classified as having this type (7.73%, Table 1). Of these patients, 21 had undergone anterior fusion to L3, 9 had fusion to L4, and 3 had undergone anterior release and posterior fusion to L5 because of a rigid tilt of L4. For type Ic, anterior correction and fusion will achieve better results compared with long segmental posterior fixation to preserve more mobile segments. Usually an end-to-end fusion, from the upper end vertebra to the lower end vertebra, is needed for type Ic (Figure 2, available for viewing online through ArticlePlus only).

### **PUMC Type II: Double Curves**

**Subtype IIa.** Double thoracic curves (13 patients, 3.05%) (Table 1). Both curves were fused superiorly at or below T2 and inferiorly to the stable vertebra of the lower thoracic curve.

**Subtype IIb.** Thoracic curve plus thoracolumbar/lumbar curve, the former is at least 10° higher than the latter. It is further divided into 2 subtypes: IIb1 and IIb2. Subtype IIb1 should meet all the following 4 criteria: (1) without thoracolumbar/lumbar kyphosis; (2) a Cobb angle of thoracolumbar/lumbar curve  $\leq 45^\circ$ ; (3) rotation of thoracolumbar/lumbar curve less than 2°; and (4) flexibility of thoracolumbar/lumbar curve  $\geq 70\%$ . Subtype IIb2 does not meet any of the aforementioned 4 criteria. In this group, there were 57 IIb1 and 51 IIb2 cases (Table 1). Of the 57 patients with subtype IIb1, 21 underwent posterior selective thoracic fusion using 3-D instrumentations, with a minimum one-year follow-up. Both curves of the other 36 patients were fused distally to L3 or L4, and an extra 2.9 segments were fused compared with posterior selective thoracic fusion. The mean coronal Cobb angle of lumbar curve before surgery was 34.6° (25°–45°), mean flexibility was 84.7% (range 62.5% to 100%), apical rotation was within grade I (Nash-Moe method), and apical translation was 10.0 mm (range 4.5–15.4). Derotation and distraction maneuver was applied in all these cases. At the final follow-up, trunk shift was 3.6 mm (range 0–12), the mean Cobb angle of the thoracic and lumbar was 18.8° and 15.9°, respectively, and no thoracolumbar kyphosis was noted (Figure 3). Thus, for type IIb1, selective thoracic fusion to the stable vertebra of the thoracic curve is recommended, while type IIb2 needs fusion of both curves.

**Subtype IIc.** Thoracic curve plus thoracolumbar/lumbar curve, the curve magnitude difference is less than 10°. By comparing the curve flexibility, it is further divided into 3 subtypes.

- IIc1: Flexibility: Thoracic curve is more than the thoracolumbar/lumbar curve; the Cobb angle of the thoracic curve on convex bending radiograph is  $\leq 25^\circ$ .
- IIc2: Flexibility: Thoracic curve is more than thoracolumbar/lumbar curve; Cobb angle of the thoracic curve on convex bending radiograph is more than 25°.
- IIc3: Flexibility: Thoracic curve is less than the thoracolumbar/lumbar curve.

For type IIc1, selective anterior fusion of the lower curve is sufficient because the upper thoracic curve is milder and more flexible, and, thus, would compensate automatically. Selection of fusion levels is similar to Ib and Ic. For type IIc2, posterior fusion of the 2 curves is recommended, but if the rotation of the lower curve is larger than grade II or the Cobb angle is more than 65°, then an anterior correction and fusion of the lower curve combined with a posterior fusion of the 2 curves are necessary. For type IIc3, selective thoracic fusion or double curve fusion should be performed based on the fusion criteria of type IIb.

**Subtype IId.** Thoracic curve plus thoracolumbar/lumbar curve, the former is 10° smaller than the latter.



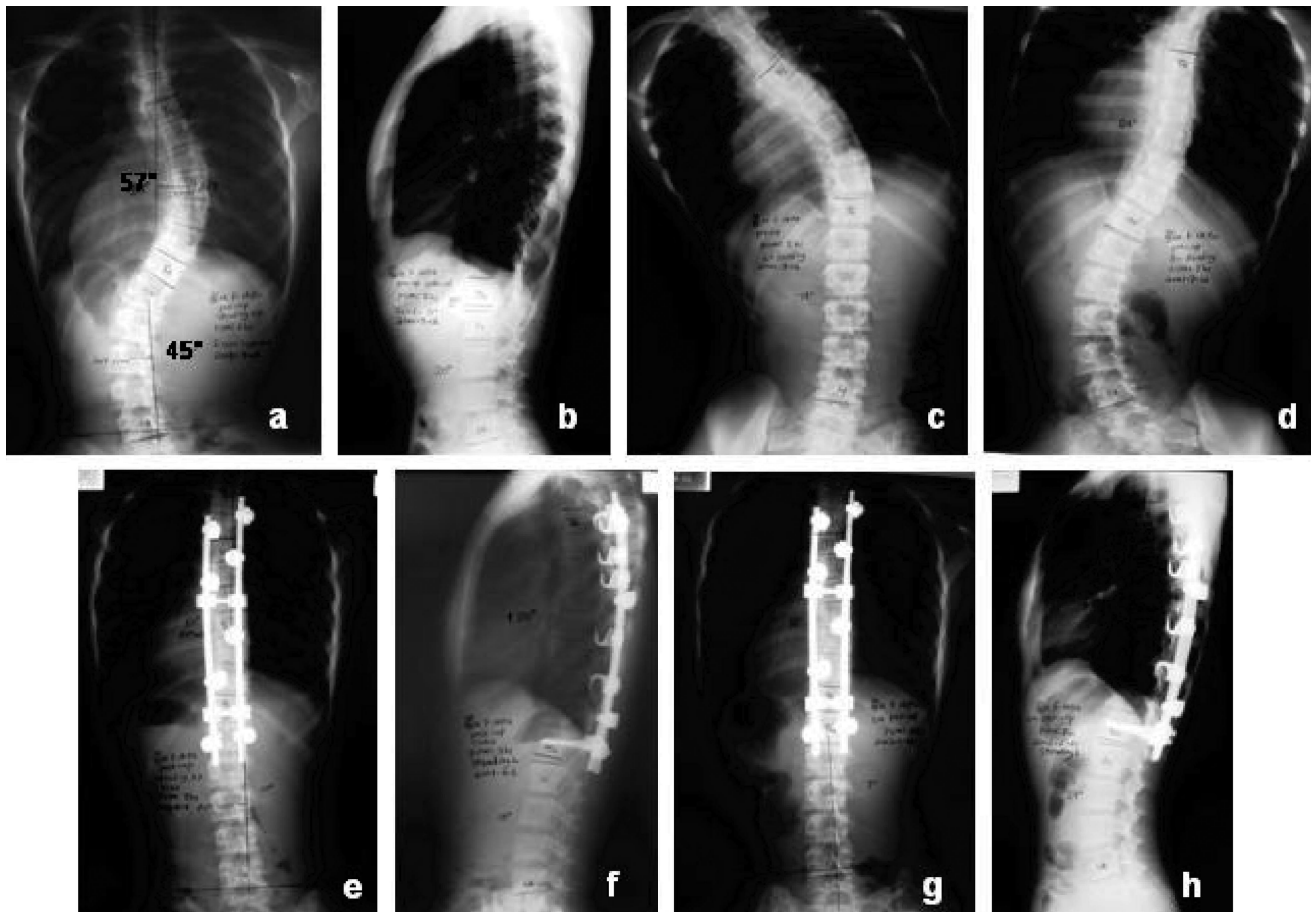


Figure 3. A 14-year-old-female with PUMC type IIb1. **a**, The thoracic curve was 57° and the lumbar curve 45°, lumbar apical vertical translation was 19 mm, while apical vertebral rotation (AVR) was grade I. **b**, No thoracolumbar kyphosis was found. **c**, **d**, The supine bending films showed the lumbar flexibility was more than 70%. **e**, **f**, After selective posterior thoracic fusion with Texas Scottish Rite Hospital (instrumentation), both coronal and sagittal balance were well maintained. **g**, **h**, The radiograph at 6-month follow-up shows that the corrected thoracic curve and kyphosis are still well maintained.

This type is divided into 2 subtypes according to the flexibility of the thoracic curve:

- IId1: Cobb angle of the thoracic curve on convex bending radiograph is  $\leq 25^\circ$
- IId2: Cobb angle of the thoracic curve on convex bending radiograph is more than  $25^\circ$

For type IId1, selective anterior fusion of the lower curve is recommended, and for IId2, double curve fusion is necessary to avoid postoperative decompensation of the thoracic curve.

In this group, all 3 patients with type IIc1 and 13 of 28 patients with IId1 (Table 1) underwent selective fusion of the lumbar or thoracolumbar using the 3-D instrumentations. There was transient deterioration of trunk shift after surgery from 19.5–23.4 mm. However, all trunk shifts recovered within 20 mm at the 6-month follow-up (Figure 4, available for viewing online through ArticlePlus only). Only one patient had  $20^\circ$  of disc angle distal to the lower end fusion vertebra as a result of over derotation and compression of the lumbar (Figure 5).

### **PUMC Type III: Triple Curves**

**Subtype IIIa.** The distal curve meets the criteria of the lumbar curve of IIb1, therefore, selective fusion of the 2 proximal curves is suggested without needing to fuse the distal lumbar curve because it is milder and more flexible.

**Subtype IIIb.** All 3 curves should be fused because the distal lumbar curve is larger and more rigid. If not fusing all the curve, decompensation will definitely occur in future (Figure 6, available for viewing online through ArticlePlus only).

### **Interobserver Reliability and Intraobserver Reproducibility of the PUMC Classification System**

The results obtained by the 6 surgeons after studying the radiographic films of 29 cases are listed in Table 2. By measuring the curve magnitudes and apical translation, determining curve apexes, recording apical rotation, and calculating the curve flexibility, the 6 reviewing physicians define the curve type. Based on their agreement on each subtype, the interobserver reliability and intraob-

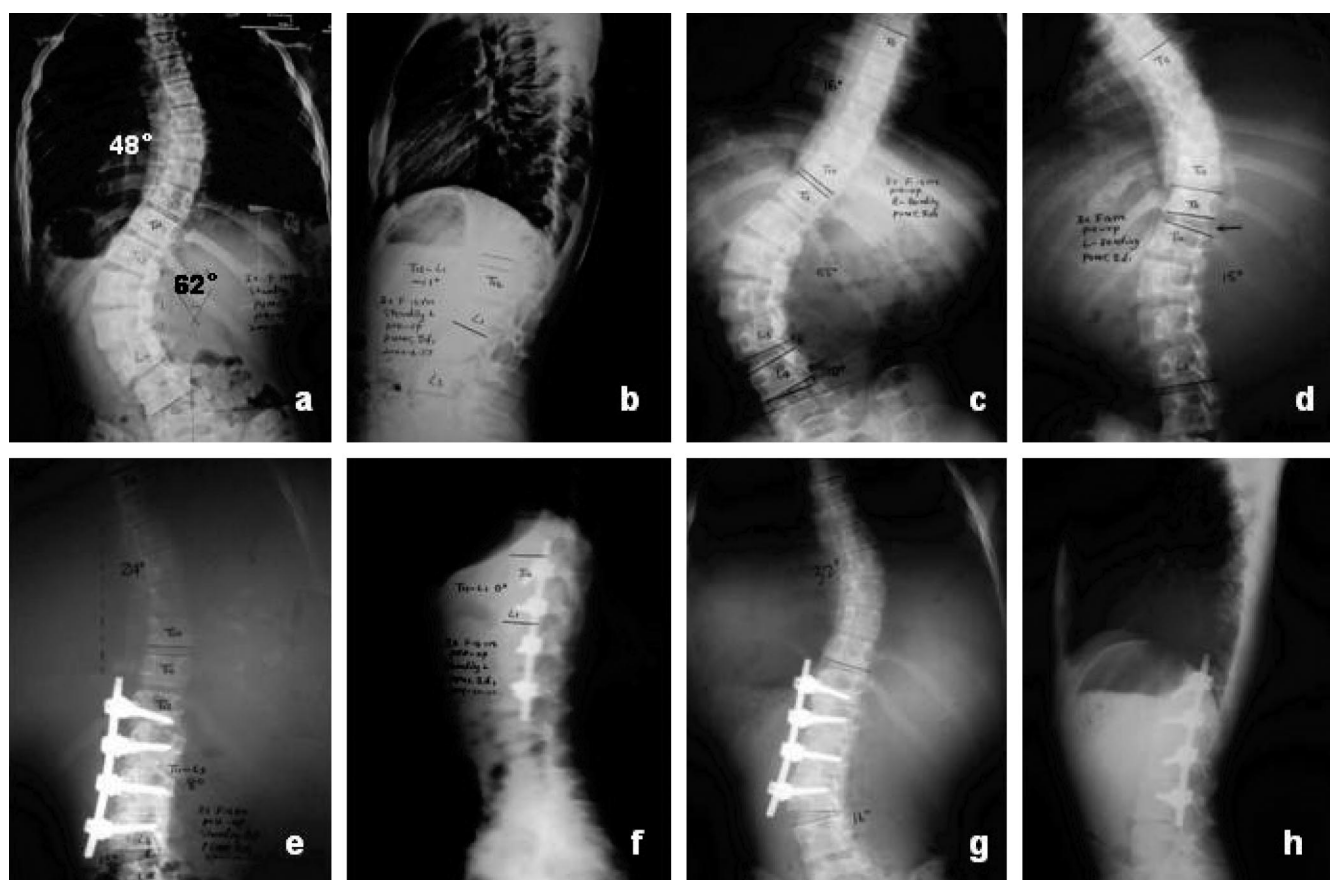


Figure 5. **a–d**, A 16-year-old female with a thoracic (48°) and lumbar curve (62°). The thoracic curve was corrected to 16° on convex bending film and was classified as type IId1. **e, f**, A selective short lumbar fusion was performed and trunk shift increased, and the disc angle at L3,4 opened because of over derotation and compression. **g, h**, At 1 1/2-year follow-up, trunk shift recovered, and disc angle was maintained at 20°.

server reproducibility were analyzed. The 6 reviewers agreed on the 13 subtypes. The more frequently, differently classified subtypes were Ia and Ib, which were regularly classified as IIb1 and Ic, respectively, because the junctional lumbar tilting is easily classified as a minor lumbar curve. The average reliability of this system was

85%, with a kappa coefficient of 0.832 ( $>0.75$ ), and mean reproducibility was 91%, with a kappa coefficient of 0.898 (Table 3).

### Clinical Validation

We have performed a prospective study of 152 cases according to this new classification system. There are 96 females and 56 males, which includes 40 cases of type Ia, 19 of Ib, 17 of Ic, 14 of IIb1, 21 of IIb2, 14 of IIc3, 16 of IId1, 7 of IId2, 2 of IIIa, and 2 of IIIb. Average age was 14.5 years (range of 10–19). Fusion was performed strictly by the prescriptive approach and fusion level. The preliminary results showed no trunk decompensation or other complications related to the selection of

**Table 2. The Analysis of the Interobserver Reliability of PUMC Classification**

Observer	No. of Same Classification	Percentage of Same Classification (N = 29)	Kappa Coefficient
1–2	25	86	0.847
1–3	26	90	0.886
1–4	24	83	0.809
1–5	24	83	0.809
1–6	25	86	0.847
2–3	24	83	0.809
2–4	23	79	0.771
2–5	24	83	0.809
2–6	23	79	0.771
3–4	25	86	0.847
3–5	25	86	0.847
3–6	26	90	0.886
4–5	26	90	0.886
4–6	25	86	0.847
5–6	24	83	0.809
Mean		85	0.832

**Table 3. The Analysis of Intraobserver Reproducibility**

Observer	Percentage of the Same Repeated Classification (N = 29)	Kappa Coefficient
1	93	0.924
2	86	0.847
3	86	0.847
4	93	0.924
5	93	0.924
6	93	0.924
Mean	91	0.898

fusion level. The case in Figure 6 was reviewed in the retrospective study, and it is not the case included in the prospective study.

## ■ Discussion

Classification of idiopathic scoliosis has always been an area of contention in spinal surgery because it has a direct relationship with the surgical outcomes. The aim of classification is to assist the selection of an appropriate approach and range of vertebral fusion levels. An ideal classification system should have the following characteristics:

1. Comprehensiveness: All types of common idiopathic scoliosis should be considered easily, not only the coronal and sagittal deformities but also the axial deformities.
2. It should be easily understood and remembered, and possess a level of reliability and reproducibility.
3. Each type of curve should correspond to an appropriate surgical procedure and fusion level to guide the treatment selection process.

Schulthess<sup>3a</sup> was the first to classify idiopathic scoliosis using the categories cervicothoracic, thoracic, thoracolumbar, lumbar, and double curves. Thereafter, other investigators found that cervicothoracic curve was nonexistent and classified idiopathic scoliosis into between 5 and 9 types.<sup>5</sup> Winter and Lonstein<sup>6</sup> divided it into 7 types according to the morphologic pattern of the curvature. Coonrad *et al*<sup>7</sup> classified it into 9 types, and 11 subtypes in terms of the location of the apex and direction of the curve after analysis of 2000 cases of idiopathic scoliosis. They provided a morphologic database of idiopathic scoliosis, which can help analyze the distribution of different kinds of scoliosis.

King *et al*<sup>1</sup> first described 5 type thoracic curves and their related fusion levels. It is the first operative classification system in history, and many surgeons used it as a guideline in their practice. However, this system only included the thoracic curves and classified them only in the coronal plane, so many problems occurred when it was applied for the treatment of idiopathic scoliosis using 3-D instrumentation. The most common problem is postoperative decompensation after selective thoracic fusion in King type II curves,<sup>13–16</sup> and comparative studies among different institutes and procedures are limited because of its lower reliability (64%) and reproducibility (69%).<sup>17,18</sup>

Lenke *et al*<sup>19</sup> developed a new system that included an analysis of coronal and sagittal deformities. It contains 6 types with 1–9 subtypes in each. It is a relatively comprehensive classification system, with a high reliability rate of 92% and reproducibility rate of 83%, which are much higher than the King-Moe system. However, the description of curvatures in the Lenke system can be misunderstood because it involves a more complex concept of structural curves, which is still controversial.

Ogon *et al*<sup>20</sup> reported that the reliability of the Lenke system was only 41%, which was much lower than that reported in the original study by Lenke *et al*<sup>19</sup>. In addition, it does not give a guide to appropriate fusion levels and surgical approach for each type of curve.

With the financial support of the Ministry of Health of the People's Republic of China, the PUMC Classification System was established based on 20 years of patient follow-up, data collection, and analysis of the large number of scoliosis cases in our hospital. It classifies idiopathic scoliosis into 3 types according to the number of curves, the SRS definition of curvature, and apex. Each type is divided into several subtypes depending on the characteristics of 3-D deformities and the flexibility of the curvature. The PUMC classification system recognizes all coronal, sagittal, and also axial spinal deformities. Therefore, it considers a wider range of patients with scoliosis, including those with delayed diagnosis or very severe and rigid deformities.

The most outstanding characteristic of this system is that with each classification type, the appropriate surgical approach and fusion level are provided, maintaining as many mobile segments as possible. The most challenging part of this system seems to be type II, which looks much too complicated. However, it is easier to be understood and remembered if the general principle of fusing as few segments as possible is realized, and that is why it has relatively higher reliability and reproducibility, 85% and 91%, respectively. Every classification system seemed to have much lower reliability and reproducibility at other centers<sup>17,19,20</sup> than the original study. Therefore, further study of the reliability and reproducibility of this system at other spinal centers is needed. Also, a long-term prospective study is mandatory to modify and improve this system.

## ■ Key Points

- The PUMC classification system categorizes idiopathic scoliosis into 3 types according to the number of curves, the SRS definition of curvature, and apex.
- Each type is divided into several subtypes depending on the characteristics of 3-D deformities and the flexibility of the curvature.
- The classification system recognizes all coronal, sagittal, and also axial spinal deformities, and is useful for selecting a surgical approach and fusion level.
- It is also easily understood and remembered, and has better interobserver reliability and intraobserver reproducibility than the King system.

## References

1. King HA, Moe JH, Bradford DS, et al. The selection of fusion levels in thoracic idiopathic scoliosis. *J Bone Joint Surg Am* 1983;65:1302–13.
2. Klockner C, Walter G, Matussek J, et al. Ventrodorsal correction and instrumentation in idiopathic scoliosis *Orthopade* 2000;29:571–7.



3. Moore MR, Baynham GC, Brown CW, et al. Analysis of factors related to trunclal decompensation following Cotrel-Dubousset instrumentation. *J Spinal Disord* 1991;4:188–92.
- 3a. Schulthess W. Die pathologie and therapie der ruckgrats-verkrummungen in Joachimsthal-Hand-Buch der Orthopadischen Chirurgie, Bd. 1a Pt. 2. Jena: Gustav Fischer, 1905–1907.
4. Richards BS. Lumbar curve response in type 2 idiopathic scoliosis after posterior instrumentation of the thoracic curve. *Spine* 1992;17:S282–6.
5. Travaglini F. Multiple primary idiopathic scoliosis. *Ital J Orthop Traumatol* 1975; 1167–80.
6. Winter RB, Lonstein JE. Idiopathic scoliosis. In: Rothman RH, Simeone FA, eds. *The Spine*. 3rd ed. Philadelphia, PA: Saunders; 1992:373–430.
7. Coonrad RW, Murrell GAC, Motley G, et al. A logical Coronal Pattern Classification of 2000 Consecutive idiopathic scoliosis cases based on the Scoliosis Research Society-defined apical Vertebra. *Spine* 1998;23:1380–90.
8. Terminology Committee of the Scoliosis Research Society. *A Glossary of Definitions. SRS Terminology Committee Glossary*. 1981, 1984, 1988.
9. Morrissy RT, Goldsmith GS, Hall EC, et al. Measurement of the Cobb angle on radiographs of patients who have scoliosis. Evaluation of intrinsic error. *J Bone Joint Surg Am* 1990;72:320–7.
10. Loder RT, Urquhart A, Graziano G, et al. Variability in Cobb angle measurements in children with congenital scoliosis. *J Bone Joint Surg Br* 1995; 77:768–70.
11. Nash CL Jr, Moe JH. A study of vertebral rotation. *J Bone Joint Surg Am* 1969;51:223–9.
12. Hall JE, Millis MB, Snyder BD. Short segment anterior instrumentation for thoracolumbar scoliosis. In: Bridwell KH, DeWald RL, eds. *The Textbook of Surgery*. 2nd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 1997: 665–74.
13. McCance S, Denis F, Lonstein J, et al. Coronal and sagittal balance in surgically treated adolescent idiopathic scoliosis with the King II curve pattern. *Spine* 1998;23:2063–73.
14. Ibrahim K, Benson L. Cotrel-Dubousset instrumentation for double major ‘right thoracic left lumbar’ scoliosis, the relation between frontal balance, hook configuration, and fusion levels [abstract]. *Orthop Trans* 1991;15:114.
15. Roye DP, Farcy JP, Rickert JB, et al. Results of spinal instrumentation of adolescent idiopathic scoliosis by King type. *Spine* 1992;17:S270–3.
16. Knapp DR, Price CT, Jones ET, et al. Choosing fusion levels in progressive thoracic idiopathic scoliosis. *Spine* 1992;17:1159–65.
17. Cummings RJ, Loveless EA, Campbell J, et al. Inter-observer reliability and intra-observer reproducibility of the system of King et al. for the classification of adolescent idiopathic scoliosis. *J Bone Joint Surg Am* 1998;80:1107–11.
18. Lenke LG, Betz RR, Bridwell KH, et al. Intra-observer reliability of the classification of thoracic adolescent idiopathic scoliosis. *J Bone Joint Surg Am* 1998;80:1097–106.
19. Lenke LG, Betz RR, Harms J, et al. Adolescent idiopathic scoliosis. A new classification to determine extent of spinal arthrodesis. *J Bone Joint Surg Am* 2001;83:1169–81.
20. Ogon M, Giesinger K, Behensky H, et al. Inter-observer and intra-observer reliability of Lenke’s new scoliosis classification system. *Spine* 2002;27:858–62.

# A Blinded Assessment of Radiographic Criteria for Atlanto-occipital Dislocation

Kristine Dziurzynski, MD, MSPH,\* Paul A. Anderson, MD,\*† Darren B. Bean, MD,‡  
James Choi, MD,§ Glen E. Levenson, PhD,¶ Rigoberto L. Marin, BS,\*\* and  
Daniel K. Resnick, MD\*

**Study Design.** Blinded comparison of 5 methods to diagnose atlanto-occipital dislocation (AOD) on plain radiographs and computerized tomography (CT) of the cervical spine.

**Objective.** To determine the best method to diagnose AOD.

**Summary of Background Data.** Several methods are proposed for the diagnosis of AOD, including the Power's ratio, X-line method, basion-dens interval, condylar gap, and Harris method. No blinded comparison of the results of these methods has been compared to patient outcome, and there is no information available regarding the accuracy of these methods applied to CT scans.

**Methods.** Plain lateral radiographs and CTs of the cervical spine were reviewed in 104 patients, including 6 with AOD. Images underwent a blinded review by a board certified neurosurgeon (D.K.R.), orthopedist (P.A.A.), radiologist (J.C.), and emergency physician (D.B.B.). Each diagnostic method for AOD was applied for determination of sensitivity, specificity, and positive and negative predictive values. The ability to identify relevant anatomic landmarks was also tabulated.

**Results.** Average values for sensitivities, specificities, positive and negative predictive values for each method applied to plain radiographs are: 0.4625–1.0, 0.8933–0.9725, 0.2775–0.45, and 0.975–1.0, respectively. These values for each method applied to CT scans are: 0.7075–1.0, 0.8725–0.9775, 0.3175–1, and 0.98–1.0, respectively. Identification of relevant anatomic landmarks occurred 99.75% of the time when these methods were applied to CT scans compared to 39% to 84% of the time on plain radiographs.

**Conclusions.** Sensitivity, specificity, positive and negative predictive values of these methods improve when applied to CT scans because of better visualization of anatomic landmarks. This result suggests CT scans of the cervical spine may be warranted in all trauma patients suspected of having cervical spine injury.

**Key words:** atlanto-occipital dislocation, Power's ratio, X-line method, basion-dens interval, condylar gap, Harris method, cervical spine injury. **Spine 2005;30:1427–1432**

Historically, atlanto-occipital dislocation (AOD) has been a difficult injury to identify on plain lateral cervical radiographs because the anatomic structures involved are poorly visualized on the film. This poor visualization may lead to missed injuries. *The Guidelines for the Management of Acute Cervical Spine and Spinal Cord Injuries* (Neurosurgery supplement, 2002) recommends applying the basion-axial interval-basion-dens interval (Harris method) to a plain lateral cervical radiograph. In the event of a nondiagnostic film in the presence of prevertebral soft tissue swelling or clinical suspicion, additional imaging with computerized tomography (CT) or magnetic resonance is recommended. These recommendations are based on only class III evidence. Because there is widespread variation among reported sensitivities of techniques to diagnose AOD,<sup>1–3</sup> we performed a blinded review of plain lateral radiographs and CT scans of the cervical spine to compare the diagnostic accuracy of the Power's ratio, X-line method, basion-dens interval, and basion-axial interval-basion-dens interval (BDI). This study was conducted to determine the best diagnostic method combined with imaging modality to identify AOD.

## ■ Materials and Methods

This is a blinded review of the plain lateral radiographs and CT scans of the cervical spine in 104 patients. Six patients had clinically confirmed AOD, and 98 were found not to have AOD based on imaging studies and subsequent clinical follow-up. Pediatric (children younger than 12 years) and adult patients were included in the review. All patients were evaluated in the emergency department of a major university hospital that is also a level one trauma center.

A board certified neurosurgeon (D.K.R.), orthopedic surgeon (P.A.A.), radiologist (J.C.), and emergency physician (D.B.B.) reviewed each image and applied the Power's ratio, X-line method, basion-dens interval, condylar gap, and basion-axial basion-dens interval (Harris method), and recorded their results (Figures 1–3). Instructions of how to apply each method was provided to the reviewers. The Power's ratio is the ratio of the basion-posterior atlas arch to the opisthion-anterior atlas arch and is abnormal at values more than one (Figures 1A, 2A, 3A).<sup>4</sup> The X-line method is considered abnormal if the line from the basion to the axis spinolaminar junction does not intersect C2, and a line from the opisthion to the posterior inferior corner of the body of the axis does not intersect C1 (Figures 1B, 2B, 3B).<sup>3</sup> The basion-dens interval is considered abnormal in the presence of a displacement >10 (12 mm in pediatrics) mm between these 2 structures (Figures 1C, 2C, 3C).<sup>5</sup> A distance >2 (5 mm in pediatrics) mm between the

From the Departments of \*Neurological Surgery, †Orthopedic Surgery and Rehabilitation, ‡Medicine, Section of Emergency Medicine, §Radiology, and ¶Surgery, University of Wisconsin-Madison; and the \*\*University of Wisconsin Medical School, Madison, WI.

Acknowledgment date: June 2, 2004. First revision date: July 17, 2004. Acceptance date: July 20, 2004.

The manuscript submitted does not contain information about medical device(s)/drug(s).

No funds were received in support of this work. No benefits in any form have been or will be received from a commercial party related directly or indirectly to the subject of this manuscript.

Address correspondence and requests for reprints to Kristine Dziurzynski, MD, MSPH, Department of Neurological Surgery, University of Wisconsin-Madison, 600 Highland Avenue, Box 8660, Madison, WI 53792; E-mail: dziurzynski@neurosurg.wisc.edu

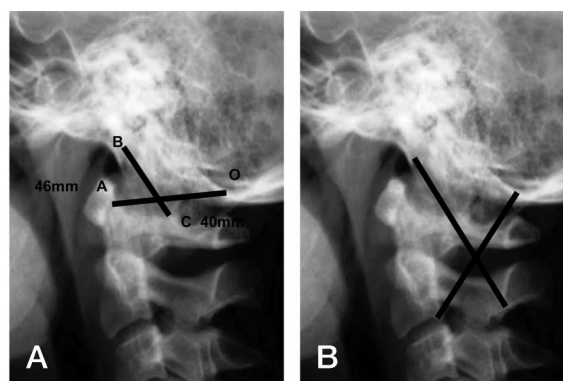
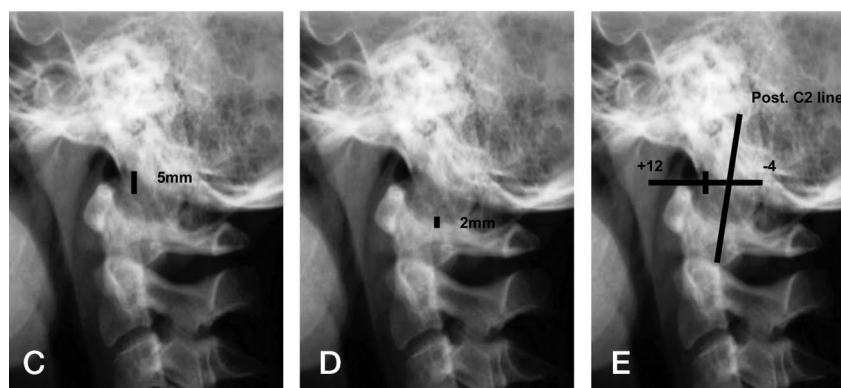


Figure 1. Methods to diagnose AOD shown on plain radiographs of a patient without injury. **A**, Power's ratio-line segments BC/OA, 40/46 mm = 0.87. **B**, X-line method. Posterior-superior aspect of dens and spinolaminar line at C1 are intersected. **C**, Basion-dens interval <10 mm. **D**, Condylar gap, 2 mm with no displacement. **E**, Harris method. Basion-dens interval of 5 mm with no displacement.



occipital condyle and the superior articular facet of the atlas is considered abnormal for the condylar gap method (Figures 1D, 2D, 3D).<sup>6</sup> Finally, anterior displacement >+12 mm or posterior displacement >-4 mm between the basion and posterior C2 line, or displacement >12 mm from the basion to the dens is considered abnormal by the basion-axial basion-dens inter-

val (Harris method) (Figures 1E, 2E, 3E).<sup>2</sup> These methods were applied to a plain lateral radiograph and midsagittal image, or parasagittal image in the case of condylar gap, of a 2-dimensional (2-D) reconstructed CT scans of the cervical spine.

Radiographs were stored as Digital Imaging and Communications in Medicine files on compact discs containing plain

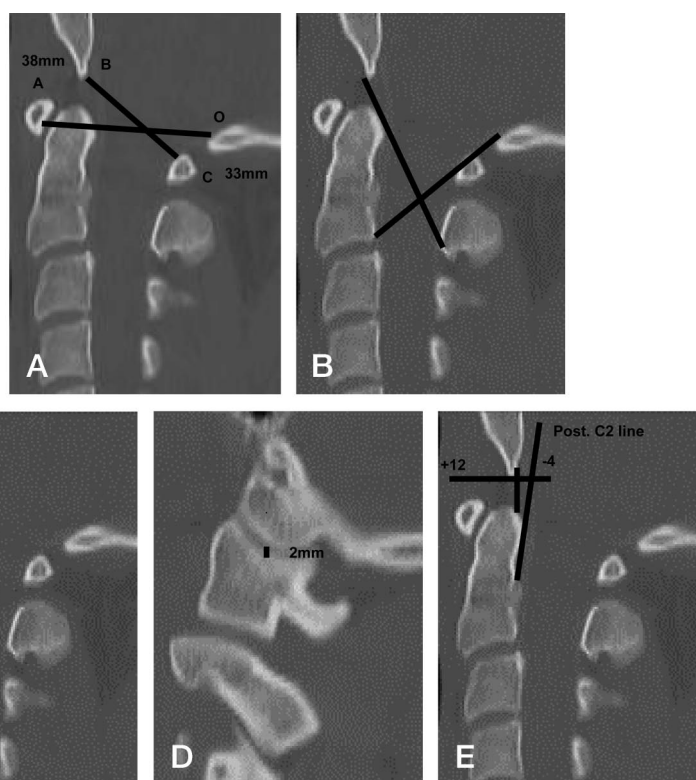


Figure 2. Methods to diagnose AOD shown on CT scans of a patient without injury. **A**, Power's ratio-line segments BC/OA, 33/38 mm = 0.87. **B**, X-line method. Posterior-superior aspect of dens and spinolaminar line at C1 are intersected. **C**, Basion-dens interval, <10 mm. **D**, Condylar gap, 2 mm with no displacement. **E**, Harris method. Basion-dens interval of 7 mm with no displacement.



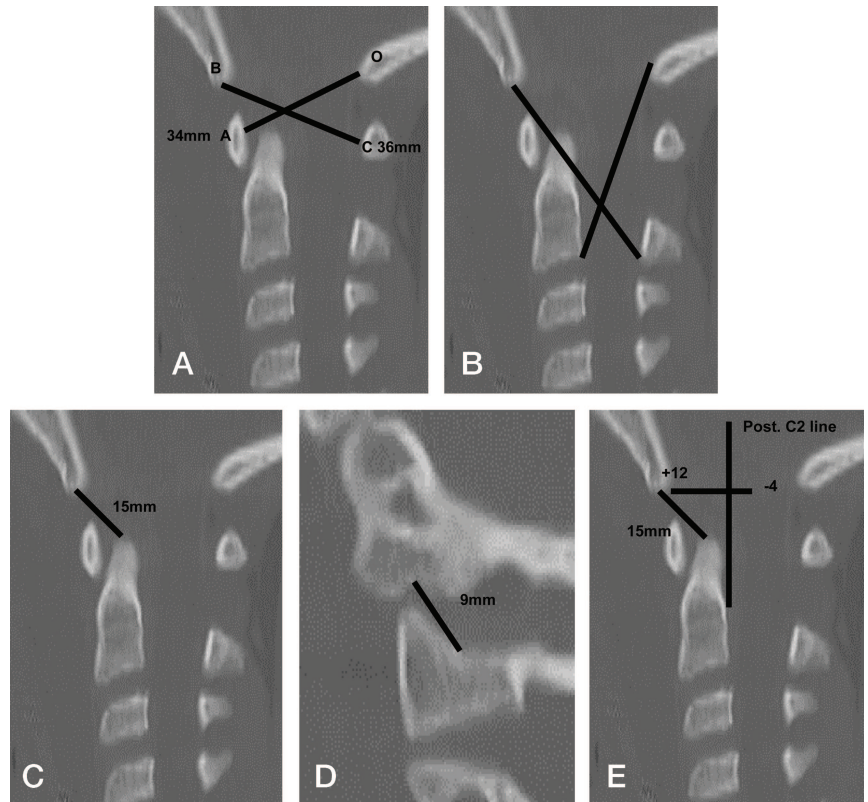


Figure 3. AOD shown on the CT scans of a pediatric patient. **A**, Power's ratio-line segments BC/OA, 36/34 mm = 1.1. **B**, X-line method. Posterior-superior aspect of dens and spinolaminar line at C1 are not intersected. **C**, Basion-dens interval, 15 mm. **D**, Condylar gap is 9 mm, 4 mm higher than the maximum 5 mm allowed in pediatric patients. It is anteriorly displaced. **E**, Harris method. Basion-dens interval of 15 mm, higher than 12 mm allowed for this method and anteriorly displaced.

radiographic and CT images of the cervical spine. CT images included axial, sagittal, and coronal reconstructions. Because every image was digitized, we were able to remove all patient identifying data and dates while generating discs. Each compact disc contained a software program, eFilm™ Lite™ (eFilm Medical Inc., Milwaukee, WI), that read the images and would automatically load onto a reviewer's computer. Although a reviewer could control image quality and magnification with built-in commands on the image-reading program, resolution was limited to that of a standard personal computer. On 3 occasions, an original film with patient identifiers removed was provided to the reviewers because of software problems reading the digitized image.

Discs were distributed to reviewers along with instructions of how to apply each method and datasheets to record their results. Data collected included calculated values for the Power's ratio, basion-dens interval, condylar gap, and yes/no answers to appropriate points of intersection for the X-line and Harris methods. Additionally, we collected data regarding how often reviewers were able to identify anatomic landmarks.

For each technique we calculated the sensitivity, specificity, positive and negative predictive values, interobserver variability, and percentage of time landmarks were visualized on films. The results were averaged for all reviewers. Data were analyzed using the Student *t* test with a value of 0.05 used to establish statistical significance. SAS statistical software (version 6.12, SAS Institute Inc., Cary, NC) was used to calculate these values. This study was approved by the University of Wisconsin-Madison Health Sciences Human Subjects Committee.

## Results

Sensitivity, specificity, positive and negative predictive values were calculated for each method. These values

were tabulated independently for each reviewer per radiographic modality and then averaged together, grouped by modality (Figures 4–7). The percentage of time reviewers were able to identify relevant landmarks on the images is presented in Figure 8. These results are presented here because of their effect on the calculation of sensitivities, specificities, positive and negative predictive values. If landmarks could not be identified on the image, the method being reviewed could not be applied. Hence, there were no results for that image to include in the calculation. Using the midsagittal view, or parasagittal for the condylar gap, on a 2-D reconstructed CT scans, relevant anatomic landmarks were identified 99.75% to 100% of the time (Figure 8).

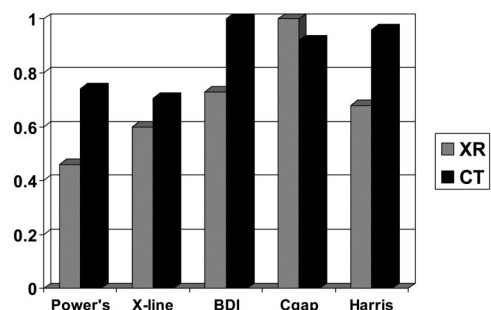


Figure 4. Average sensitivities for the 5 methods studied. Power's ratio x-ray (XR) (0.46), X-line XR (0.60), basion-dens interval (BDI) XR (0.73), condylar gap (Cgap) XR (1.0), and Harris method XR (0.68). Power's ratio CT scans (0.74), X-line CT scans (0.71), BDI CT scans (1.0), Cgap CT scans (0.92), and Harris method CT scans (0.96).

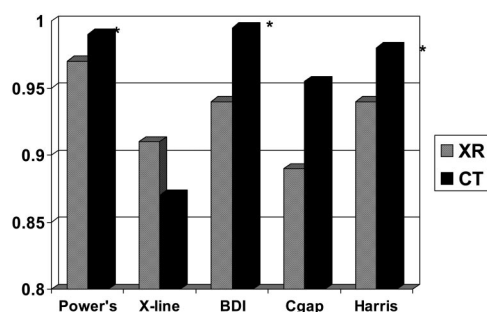


Figure 5. Average specificities for the 5 methods studied. Power's ratio x-ray (XR) (0.97), X-line XR (0.91), basion-dens interval (BDI) XR (0.94), condylar gap (Cgap) XR (0.89), and Harris method XR (0.94). Power's ratio CT scans (0.99), X-line CT scans (0.87), BDI CT scans (0.995), Cgap CT scans (0.955), and Harris method CT scans (0.98). The asterisk denotes statistical significance of the difference in specificities between XR and CT scans ( $P < 0.05$ ).

### Power's Ratio

The average sensitivity, specificity, positive and negative predictive values for the Power's ratio applied to lateral radiographs of the cervical spine are 0.46, 0.97, 0.44, and 0.98, respectively. Applied to CT scans, the results for sensitivity, specificity, positive and negative predictive values are 0.74, 0.99, 0.81, and 0.99, respectively. The differences between the values of specificity and positive predictive value between plain radiographs and CT scans were statistically significant ( $P < 0.05$ ). There was a nonsignificant trend for improvement in sensitivity ( $P = 0.08$ ).

### X-Line Method

For the X-line method applied to lateral radiographs of the cervical spine, the average sensitivity, specificity, positive and negative predictive values are 0.60, 0.91, 0.28, and 0.98, respectively. Applied to CT scans, the results for sensitivity, specificity, positive and negative predictive values are 0.71, 0.87, 0.32, and 0.98, respectively.

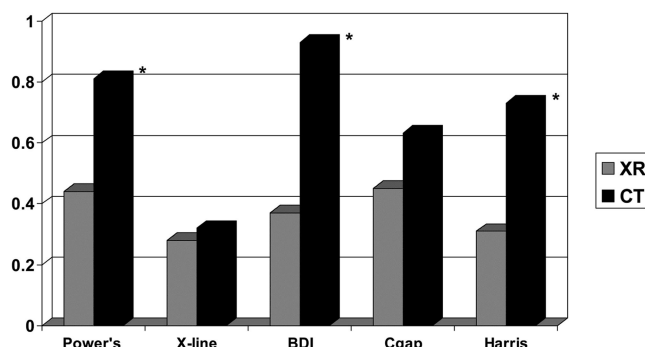


Figure 6. Average positive predictive values for the 5 methods studied. Power's ratio x-ray (XR) (0.44), X-line XR (0.28), basion-dens interval (BDI) XR (0.37), condylar gap (Cgap) XR (0.45), and Harris method XR (0.31). Power's ratio CT scans (0.81), X-line method CT scans (0.32), BDI CT scans (0.93), Cgap CT scans (0.63), and Harris method (0.73). The asterisk denotes statistical significance of the difference in positive predictive values between XR and CT scans ( $P < 0.05$ ).

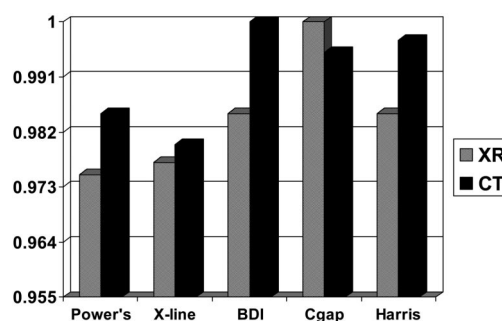


Figure 7. Average negative predictive values for the 5 methods studied. Power's ratio x-ray (XR) (0.975), X-line XR (0.977), basion-dens interval (BDI) XR (0.985), condylar gap (Cgap) XR (1.0), and Harris method XR (0.985). Power's ratio CT scans (0.985), X-line CT scans (0.98), BDI CT scans (1.0), Cgap CT scans (0.995), and Harris method CT scans (0.995).

There were no statistically significant differences among any of these values between plain radiographs and CT scans.

### Basion-Dens Interval

The average sensitivity, specificity, and positive and negative predictive values for the basion-dens interval applied to lateral radiographs of the cervical spine are 0.73, 0.94, 0.37, and 0.99, respectively. Applied to CT scans, the results for sensitivity, specificity, and positive and negative predictive values are 1.0, 0.995, 0.93, and 1.0, respectively. The differences between the values of specificity and positive predictive value between plain radiographs and CT scans were statistically significant ( $P < 0.05$ ).

### Condylar Gap Method

For the condylar gap method applied to lateral radiographs of the cervical spine, the average sensitivity, specificity, and positive and negative predictive values are 1.0, 0.89, 0.45, and 1.0, respectively. Applied to CT scans, the results for sensitivity, specificity, and positive

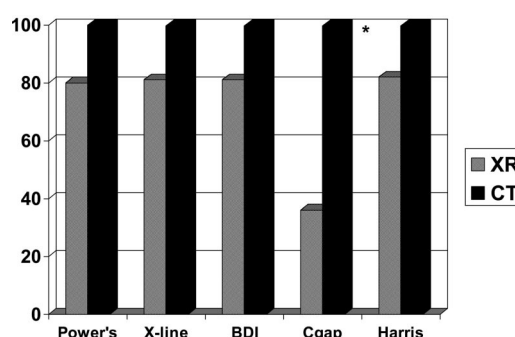


Figure 8. Average rates of readability. Percentage of the time reviewers were able to identify relevant anatomic landmarks on x-ray (XR) and CT scans. Power's ratio XR (80%), X-line XR (81%), basion-dens interval (BDI) XR (81%), condylar gap (Cgap) XR (36%), and Harris method XR (82%). Power's ratio CT scans (100%) and X-line, BDI, Cgap, Harris (99.75%). The asterisk denotes statistical significance in the difference between readability of XR and CT scans ( $P < 0.05$ ).

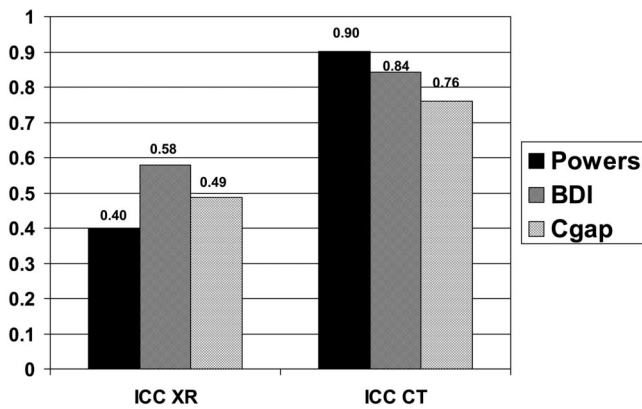


Figure 9. The intra-class correlation coefficient (ICC) is an index of reliability that can measure interobserver variability. Values closer to one represent decreased variability. Here, for the Power's ratio, basion-dens interval (BDI) and condylar gap (Cgap) applied to CT scans the values are closer to one when compared to x-ray (XR). This shows there is more variability among reviewers when these methods are applied to x-rays.

and negative predictive values are 0.92, 0.955, 0.63, and 0.995, respectively. There were no statistically significant differences among any of these values between plain radiographs and CT scans.

#### **Basion-Axial Interval-Basion-Dens-Interval (Harris method)**

The average sensitivity, specificity, positive and negative predictive values for the basion-axial interval-basion-dens-interval (Harris method) applied to lateral radiographs of the cervical spine are 0.68, 0.94, 0.31, and 0.99, respectively. Applied to CT scans, the results for sensitivity, specificity, and positive and negative predictive values are 0.96, 0.98, 0.73, and 0.998, respectively. The differences between the values of specificity and positive predictive value between plain radiographs and CT scans were statistically significant ( $P < 0.05$ ).

Except for the sensitivity of the condylar gap and specificity of the X-line method, sensitivity, specificity, and positive and negative predictive values increased when applied to the midsagittal reconstructed view on a 2-D reconstructed CT scans.

To quantify the interobserver variability, we used the intra-class correlation coefficient, a ratio of between individual variability to the sum of between individual and between reader variability. Values can range from zero to one, those closer to one represent smaller amounts of variability. This was calculated for the Power's ratio, basion-dens interval, and condylar gap methods (Figure 9). X-line and Harris methods were not included in this calculation because their determination includes nonnumerical assessments, and intra-class correlation coefficient can only be calculated for integer values. The least amount of variability was observed for the Power's ratio applied to CT scans (0.9029). Highest variability was seen in the Power's ratio applied to radiographs (0.3974). There is decreased interobserver variability among values obtained from CT scans.

## ■ Discussion

AOD is one of the most devastating injuries of the craniocervical junction. Sequelae include death, quadriplegia, paresis of one or more extremity, and cranial nerve deficits. Head injuries and fractures at other spinal levels are frequent comorbidities, confounding identification of specific pathology responsible for a patient's neurologic deficit. Rapid identification and treatment of this injury is imperative to improve survival and prevent further neurologic decline, especially in patients who present neurologically intact or have limited deficits. Therefore, a reliable, sensitive, and easy-to-apply method for the diagnosis of AOD is required.

The radiographic diagnosis of AOD may be made using plain lateral cervical radiographs by applying the Power's ratio, X-line method, basion-dens interval, condylar gap, and Harris methods. There is widespread variation in the reported accuracies of each. Depending on the study, the sensitivity of the X-line method ranges from 20% to 75%,<sup>2,3</sup> Power's ratio 33% to 60%,<sup>2,3</sup> 50% sensitivity of the basion-dens interval,<sup>3</sup> and 100% sensitivity of the Harris method. It is noteworthy that these statistics only apply to radiographs on which the required landmarks could be identified.

A universal theme underlying the difficulties in diagnosing AOD using plain lateral cervical radiographs is the ability to visualize the anatomic landmarks required for application of these methods. The basion is frequently obscured by other bony anatomy of the skull or can be "cut off" as the radiograph is shot. Both of these problems are alleviated by CT scans with 2-D reconstruction of the cervical spine. Our reviewers indicated that they were able to identify landmarks nearly 100% of the time when using CT scans. Conversely, they could only identify the condylar gap on radiographs an average of 36% of the time and measure the basion-dens-basion-axial interval (Harris method) an average of 82% of the time.

Except for the sensitivity of the condylar gap and specificity of the X-line method, the sensitivity, specificity, and positive and negative predictive values of each method improved with application to a CT scans. Sensitivities increased an average of 23 percentage points for each method when applied to CT scans, and positive predictive values increased an average of 31 percentage points. It should be noted that the condylar gap can also be measured from the open-mouthed odontoid view on a plain film. We did not include this view in our study because this image is frequently not available for patients who are likely to have AOD. Advanced Trauma Life Support guidelines recommend a lateral cervical radiograph only during the primary survey; because many patients at risk for AOD are intubated, the odontoid view is usually not obtained.

In our study, we found the basion-dens interval, defined as abnormal when  $>10$  mm, to have the highest



sensitivity (1.0), specificity (0.995), and positive (0.93) and negative predictive values (1.0) when applied to CT scans of the cervical spine. On plain radiographs, the condylar gap method has the highest sensitivity (1.0), positive predictive value (0.45), negative predictive value (1.0), while the Power's ratio has the highest specificity (0.97). These values are superior to those we obtained when we applied the Harris method, as recommended by the Guidelines for the Management of Acute Cervical Spine and Spinal Cord Injuries. Our findings support the use of the basion-dens interval (with 10 mm as the cutoff) as the diagnostic test of choice when applied to CT scans as it is the easiest to apply and appears to be the most accurate method of those evaluated. Furthermore, our findings also support the use of CT imaging of the craniocervical junction because CT scans improves the diagnostic accuracy of these methods tested, largely by improving the ability to identify anatomic landmarks.

### ■ Key Points

- AOD is a potentially fatal injury that requires rapid identification to prevent further morbidity or mortality.
- Diagnosis of AOD has historically been difficult on plain lateral cervical radiographs because of poor visibility of anatomic structures involved.

- There is better visibility of craniocervical junction anatomy on a 2-D reconstructed CT scans of the cervical spine, improving the identification of AOD.
- Sensitivities, specificities, and positive and negative predictive values of diagnostic methods for AOD improve once applied to CT scans of the cervical spine.

### Acknowledgment

The authors thank Dr. Gary J. Wendt and Brian Gruell, University of Wisconsin Hospital and Clinics, Department of Radiology, for their assistance in helping us create the compact discs.

### References

1. Section on Disorders of the Spine and Peripheral Nerves. Diagnosis and management of traumatic atlanto-occipital dislocation injuries. *Neurosurgery* 2002; 50, S105–13.
2. Harris JH Jr, Carson GC, Wanger LK, et al. Radiologic diagnosis of traumatic occipitovertebral dissociation: 2. Comparison of three methods of detecting occipitovertebral relationships on lateral radiographs of supine subjects. *AJR Am J Roentgenol* 1994;162:887–92.
3. Lee C, Woodring JH, Goldstein SJ, et al. Evaluation of traumatic atlanto-occipital dislocations. *AJNR Am J Neuroradiol* 1987;8:19–26.
4. Powers B, Miller MD, Kramer RS, et al. Traumatic anterior atlanto-occipital dislocation. *Neurosurgery* 1979;4:12–7.
5. Whooley MH, Bruwer AJ, Baker HL. The lateral roentgenogram of the neck; with comments on the atlanto-odontoid-basion relationship. *Radiology* 1958; 71:350–6.
6. Werne S. Studies in spontaneous atlas dislocations. *Acta Orthop Scan* 1957; 23:1–150.

## Classification in Nonspecific Low Back Pain: What Methods do Primary Care Clinicians Currently Use?

Peter Kent, PT, GradDip(ManipPhysio) and Jennifer L. Keating, PT, PhD

### **Study Design.** Postal survey.

**Objectives.** To describe the signs and symptoms that clinicians think represent nonspecific low back pain (NSLBP) subgroups, and to report the labels that clinicians give to those subgroups.

**Summary of Background Data.** The cause of most low back pain (LBP) cannot be diagnosed. Consequently, approximately 80% of primary care LBP presentations are most accurately labeled as NSLBP. Most Australian primary care clinicians think that NSLBP is heterogeneous and treat patients differently based on that heterogeneity. This research sought to identify the subgroups clinicians believe are recognizable within that heterogeneity.

**Methods.** Analysis of survey data from 651 primary care clinicians from 6 professional disciplines: physiotherapy, manipulative physiotherapy, chiropractic, osteopathy, general medicine, and musculoskeletal medicine.

**Results.** There was little consensus among participating clinicians regarding the signs and symptoms that identify NSLBP subgroups. Most clinicians give labels to NSLBP subgroups that imply putative pathoanatomy, however, the evidence that these labels are valid is scant and controversial.

**Conclusions.** A lack of consensus among participating clinicians regarding NSLBP subgroups and a lack of evidence for the validity of NSLBP subgrouping are a compelling argument for further research into this clinical practice.

**Key words:** nonspecific low back pain, treatment, classification, diagnosis, subgroups. *Spine* 2005;30:1433–1440

Nonspecific low back pain (NSLBP) represents approximately 80% of primary care low back pain (LBP) presentations.<sup>1,2</sup> Currently, it is not possible for clinicians to reach NSLBP pathoanatomical diagnoses with certainty because of the poor correlation between our knowledge of pathoanatomy and clinical presentation. In the absence of diagnostic precision, most Australian primary care clinicians use signs and symptoms to classify patients with NSLBP.<sup>3</sup> They do this because most primary

care clinicians believe that NSLBP is heterogeneous and treat patients with NSLBP differently based on that heterogeneity.<sup>3</sup> This practice is prevalent even though there is little evidence for its validity, and despite the arguments of some investigators that NSLBP outcomes would be improved if NSLBP were considered homogeneous and treated with a generic approach.<sup>4–9</sup>

There are many advocates of NSLBP heterogeneity and many classification schemes that have been suggested.<sup>10–24</sup> Advocates of NSLBP heterogeneity argue that recognizing patterns of signs and symptoms as identifiers of discrete subgroups assists with making a prognosis and allows refinement of treatment selection.

Currently, there is a paucity of evidence that proposed NSLBP subgrouping schemes validly infer different prognoses. There are known risk factors for NSLBP chronicity,<sup>25–40</sup> but these factors are not typically included in subgrouping schemes. We are not aware of any longitudinal outcomes data that test the prognostic use of specific NSLBP subgrouping schemes. There are studies that show the potential prognostic use of isolated assessment findings such as centralization<sup>41–45</sup> but not the prognostic use of the combined components of a comprehensive subgrouping scheme.

We propose that NSLBP subgrouping only has treatment selection use if this practice positively impacts patient outcome. Currently, the evidence for this is sparse and of varying quality. Only a small number of randomized controlled trials have quantified the therapeutic effects achieved for patients treated with the protocols suggested by specific subgrouping schemes, compared with the therapeutic effects achieved with generic treatment protocols. Some trials have shown higher therapeutic effect with NSLBP subgrouping protocols,<sup>46–49</sup> and some have not.<sup>50,51</sup> Randomized controlled trials showing higher therapeutic effects with subgrouping protocols included differential treatment based on those protocols. Hence, when outcomes differ for subgroups, it is not clear whether it is because of the different treatment or the membership of a particular subgroup. What is not known is whether the prognosis for subgroups varies when no treatment is provided or when treatment is uniform.

Given the prevalence of NSLBP, its social and economic costs, and the prevalence of subgrouping in NSLBP treatment, there is a compelling argument for further longitudinal studies that examine the relationship between subgroup membership and outcome. A preliminary step in this process is knowledge of the current subgrouping practices of primary care clinicians. Studies in Britain, Ireland, and Washington state report that

From the School of Physiotherapy, La Trobe University, Melbourne, Victoria, Australia.

Acknowledgment date: December 8, 2003. First revision date: April 18, 2004. Second revision date: July 27, 2004. Acceptance date: August 2, 2004.

Supported by Faculty of Health Sciences (La Trobe University), Joint Coal Board Health & Safety Trust (Australia), and Musculoskeletal Physiotherapy Association (Victoria).

The manuscript submitted does not contain information about medical device(s)/drug(s).

Corporate/Industry, Federal, Institutional, Foundation, and Professional Organizational funds were received in support of this work. No benefits in any form have been or will be received from a commercial party related directly or indirectly to the subject of this manuscript.

Address correspondence and reprint requests to Peter Kent, PT, GradDip(ManipPhysio), School of Physiotherapy, La Trobe University, Bundoora, 3083, Victoria, Australia; E-mail: peter.kent@latrobe.edu.au

physiotherapists use the McKenzie subgrouping method, with 44% to 67% of patients with LBP.<sup>52-54</sup> Little is known about the physiotherapist's use of other subgrouping methods or the current subgrouping practices of clinicians across primary care professional disciplines.

Of Australian primary care clinicians who responded to a recent survey,<sup>3</sup> 74% thought that it is currently possible to recognize NSLBP subgroups, and 93% treat NSLBP differently based on patterns of signs and symptoms. Therefore, the first aim of this research was to gather data on the signs and symptoms that clinicians think represent various NSLBP subgroups. The second aim was to report the labels that clinicians give to those subgroups.

## ■ Methods

**Survey.** A survey of Australian primary care clinicians was conducted to assess aspects of their current NSLBP subgrouping practices. Questionnaires were mailed to 200 randomly selected clinicians from each of 6 primary care disciplines: chiropractors, general medical practitioners, manipulative physiotherapists, musculoskeletal medicine practitioners, osteopaths, and physiotherapists. There were 651 (59.8%) completed questionnaires returned. Details of the survey design, method, response, and the views of clinicians regarding NSLBP heterogeneity have been published.<sup>3</sup> The survey also collected data on the assessments that clinicians use in acute NSLBP, and these will be reported in a subsequent article.

The survey included 2 questions. Question 1: "There is a view in the medical and research community that LBP that is not due to serious pathology (cancer, fracture, infection etc) and not due to nerve root or thecal irritation, is best labeled 'non-specific low back pain' (previously called mechanical LBP). Do you think that it is possible for you to recognize sub-groups of non-specific low back pain and therefore reach a more precise working diagnosis (disc lesions, instability, postural syndrome, facet syndrome, etc)? Yes/No."

Question 5: "If Yes, for the three most common non-specific LBP conditions that you see, list the examination findings that you think are discriminative. For example list the findings that suggest to you that a condition is present, and (if possible) what diagnostic or descriptive label you give each of these conditions."

No temporal characteristics of NSLBP were nominated because information was sought on NSLBP clinical reasoning regardless of duration of the condition. Clusters of signs and symptoms responses to Question 5 were tabulated in a spreadsheet and reviewed. The primary researcher developed "themes" from the responses (Appendix 1, available for viewing on ArticlePlus only). For example, "limited forward bending," "limited flexion," "flexion restriction," "reduced forward bend," "unable to touch toes," and "unable to forward bend" were all listed as occurrences of the theme named "flexion reduced." Ultimately, 78 themes encapsulated all of the signs and symptoms reported by the clinicians. There were 1068 subgroup responses to Question 5 from 416 clinicians (a mean of 2.5 subgroups nominated by each responding clinician). The mean number of signs and symptoms for each subgroup response was 3.5 (3722 signs and symptoms responses in total). Two independent reviewers coded these signs and symptoms responses. Every response was successfully coded under one of these 78 themes. The initial coding resulted in 73%

complete agreement between the 2 reviewers, and the differences in initial coding (27%) were commonly of a minor nature and were resolved by negotiation.

The descriptive labels given to subgroups were also themed using the same process. For example, all labels that referred to a facet or zygapophysial joint lesion were coded as "facet syndrome." The coding labels were facet, contained disc, instability, sacroiliac joint, postural, muscle, piriformis, spondylolisthesis, degeneration, iliolumbar, miscellaneous pathoanatomy, nonpathoanatomy, and no label. The initial coding resulted in 93% complete agreement between the 2 reviewers. The coding instructions, coding categories, and precoded and coded data are available by request from the first author (P.K.).

**Data Analysis.** These subgroup data were analyzed in a number of ways. The frequencies of subgroup label use (proportions  $\pm$  95% confidence intervals) were determined for the cohort as a whole and also compared across disciplines.

The frequency with which particular signs and symptoms were nominated for particular subgroups was calculated. Arbitrarily, we only included subgroups in this analysis if they had been nominated by more than 5% of clinicians. Clinicians reported a large number of signs and symptoms, but some were reported infrequently. Arbitrarily, we only included signs or symptoms in this analysis if they met 2 criteria. First, they had to have been nominated by more than 10% of respondents who nominated that subgroup. Second, the lower limit of the 95% confidence interval around the proportion of clinicians nominating a particular sign or symptom had to be more than zero. The signs or symptoms for which there was no overlap between subgroups was determined by tabulating the signs and symptoms (by subgroup), and deleting any that were nominated in more than one subgroup.

The agreement between responses to the question on specific combinations of signs and symptoms was determined for each of the 5 most frequently nominated subgroups. All subgroup responses that were coded under the same label were identified. The most frequently nominated sign or symptom for that subgroup was then identified. The proportion of clinicians that agreed on the most common single sign or symptom as representing that particular subgroup was calculated. This process was repeated for the subgroup responses that included the 2 most frequently nominated signs or symptoms and again for the 3 most frequently nominated signs or symptoms for that particular subgroup. The proportion of clinicians that agreed that the most frequently nominated combinations of 2 specific signs or symptoms indicated particular subgroups was examined by professional discipline to determine whether this agreement traversed disciplines.

Cluster analysis was performed using a hierarchical approach on the entire data set of signs and symptoms. This cluster analysis ignored the descriptive labels, and explored underlying patterns of signs and symptoms that clinicians nominated as specific subgroups to test the possibility that disciplines might be using different labels for the same set of signs and symptoms. The Ward hierarchical clustering algorithm was used because the analysis was exploratory. Nonhierarchical clustering would have required the *a priori* estimation of nonrandom seed points (centroids).<sup>55</sup>

Tests for significant differences in proportions across disciplines were conducted using Bonferroni adjusted ( $P = 0.003$ ) inferential confidence intervals and presented graphically as



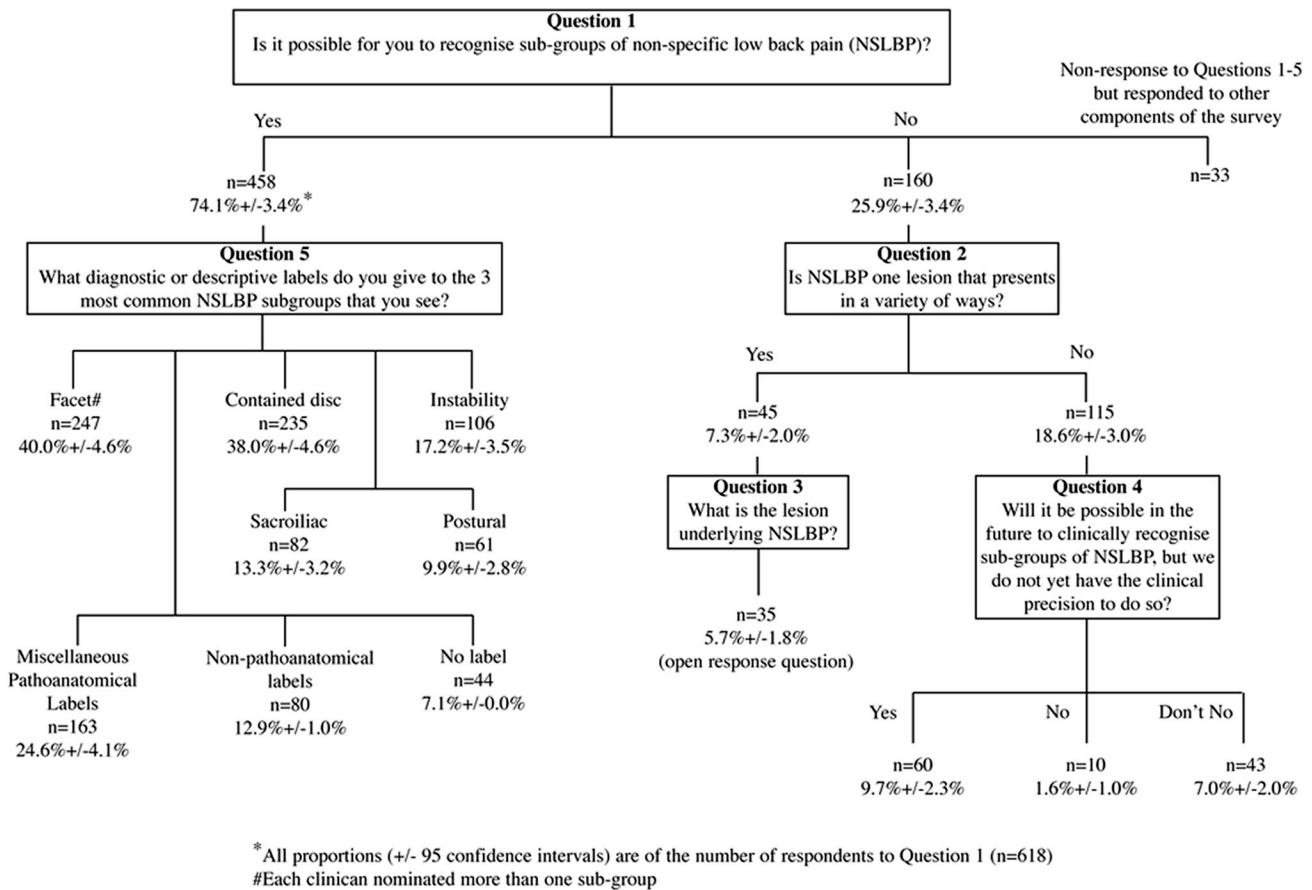


Figure 1. Clinician's responses to the survey's subgrouping questions and the proportion of clinicians nominating each subgroup.

described by Tryon.<sup>56</sup> The alpha level for each comparison was determined using the calculation: "alpha/number of comparisons" where the number of comparisons =  $(n \times (n-1))/2$ . There were 6 groups in the comparisons described, rendering 15 pair-wise comparisons:  $(n \times (n-1))/2 = (6 \times 6^{-1})/2 = 30/2 = 15$ . This resulted in resetting the alpha level for 95% confidence in results of any pair-wise comparison to  $(0.05/15) = 0.003$ . Inferential confidence intervals are Bonferroni adjusted confidence bands. Where no visual or numerical overlap occurs between these confidence bands, a difference between proportions can be concluded with 95% confidence. Statistical analysis was conducted using SPSS 10 (SPSS Inc., Chicago, IL) and Excel X (Microsoft Corp., Redmond, WA).

## Results

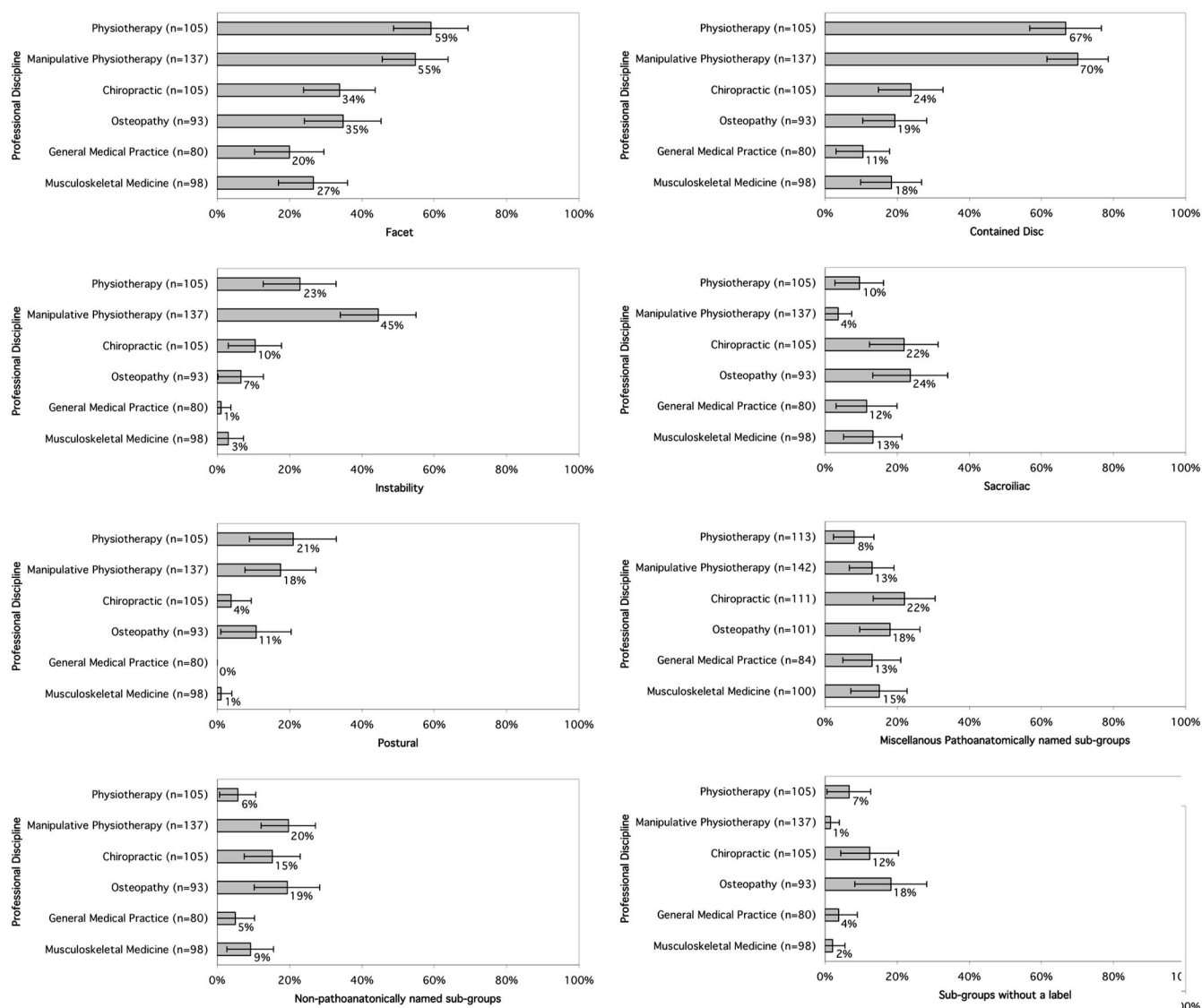
### Subgroup Labeling

The nominated subgroup labels were divided into 3 categories. The first category was referred to as "Pathoanatomical" because these labels, such as facet syndrome, contained disc lesion, imply a putative pathoanatomical source of pain. Of all nominated subgroup labels, 88.3% were in this category. The second category was referred to as "Nonpathoanatomical" because these were labels, such as acute back, nonspecific, that did not imply a pathoanatomical source of pain. Of all nominated subgroup labels, 7.5% were in this category. The third cat-

egory was made up of subgroups for which the clinician did not volunteer any label. Of all nominated subgroup labels, 4.1% were in this category. The proportions of clinicians nominating subgroups in these broad categories (*i.e.*, pathoanatomical, nonpathoanatomical, and no label) were similar across professional disciplines.

There were only 5 pathoanatomically named subgroups that were reported by more than 5% of clinicians: facet syndrome, contained disc lesion, instability, sacroiliac joint problems, and postural syndrome. Therefore, the other pathoanatomically named subgroups were included in the miscellaneous pathoanatomy subgroup for subsequent analysis. The clinician's responses to the survey's subgrouping questions and the proportions of clinicians nominating each subgroup are shown in Figure 1.

In Figure 1, the proportions of clinicians nominating particular subgroup labels have been calculated using inferential confidence intervals to allow observation of significant differences in the frequency of reporting of these labels. Facet syndrome ( $40.0\% \pm 4.6\%$ ) and contained disc ( $38.0\% \pm 4.6\%$ ) were the most frequently nominated subgroups, suggesting that clinicians consider these subgroups to be prevalent. There were significant differences in the rate of reporting of particular subgroups across disciplines (Figure 2). A higher proportion of physiotherapists and manipulative physiothera-



Respondents to Question 1 included both those who thought that it is, and those that thought it is not, currently possible to recognise NSLBP sub-groups.

The width of the 95% confidence intervals have been adjusted (Tryon, 2001) such that where no visual overlap occurs between confidence intervals for any particular comparison, a difference between these proportions can be concluded with 95% confidence (Figure 2). The proportion of all respondents to Question 1 (*i.e.*, "Is it possible to recognize NSLBP sub-groups?") from each discipline who nominated particular subgroups.

pists reported the facet and contained disc subgroups than did the other disciplines.

More physiotherapists and manipulative physiotherapists reported the instability subgroup than general medical practitioners and musculoskeletal medicine practitioners. A higher proportion of manipulative physiotherapists reported the instability subgroup than chiropractors and osteopaths. More chiropractors and osteopaths reported a sacroiliac joint subgroup than did manipulative physiotherapists.

A higher proportion of physiotherapists and manipulative physiotherapists reported the postural subgroup than general medical practitioners and musculoskeletal medicine practitioners. More osteopaths reported the

postural subgroup than did general medical practitioners. A higher proportion of osteopaths reported unnamed subgroups than did musculoskeletal medicine practitioners and manipulative physiotherapists.

### Subgroup Composition

The signs and symptoms for which there was no overlap among subgroups were determined (Table 1). The proportion of respondents nominating these signs and symptoms ranged from 9.5% to 54.7%.

The signs and symptoms for the 5 most commonly nominated subgroups (*i.e.*, facet syndrome, contained disc lesion, instability, sacroiliac joint problems, and postural syndrome) are shown in Appendices 2.1–2.5

**Table 1. The Signs and Symptoms Nominated for Particular Subgroups for Which There was No Overlap Among Subgroups**

Facet	Contained Disc	Instability	Sacroiliac	Postural
Extension: reduced (38.8% ± 6.1%)	List (31.3% ± 5.9%)	Mid range catch (44.8% ± 9.5%)	Asymmetry SIJ (palpatory or movement) (54.7% ± 10.5%)	Postural abnormalities (42.6% ± 12.4%)
Lateral flexion: reduced (22.9% ± 5.2%)	SLR: positive (30.5% ± 5.9%)	Palpatory hypermobility (27.6% ± 8.5%)	SIJ provocation test: positive (29.1% ± 9.6%)	Pain with prolonged postures (36.1% ± 12.0%)
Palpatory hypomobility (20.0% ± 5.0%)	Pain sitting (28.8% ± 5.8%)	Recurrent pain (12.4% ± 6.3%)	Pain buttock (29.1% ± 9.6%)	Muscle tightness (21.3% ± 10.3%)
Posterior quadrant: reduced (18.8% ± 4.9%)	Impulsion: positive* (23.2% ± 5.4%)	Pain: standing (10.5% ± 5.8%)	Pain: leg (19.8% ± 8.4%)	Cause: nonspecific (11.5% ± 8.0%)
Restricted movement (16.7% ± 4.7%)	Extension helps (13.7% ± 4.4%)	Imaging findings positive (9.5% ± 5.6%)		
Pain: paracentral (11.8% ± 4.0%)				

The proportions of clinicians nominating these signs and symptoms within each subgroup are shown in parentheses.

\* Coughing, laughing, Valsalva maneuver.

SIJ = sacroiliac joint; SLR = straight leg raising.

(available for viewing online through ArticlePlus only). Overall, there was low agreement on specific combinations of signs and symptoms for each of these 5 subgroups. The range of agreement for the *single* most common sign or symptom was 42.6% ± 12.4% to 54.7% ± 10.5%. Agreement on the 2 most common signs or symptoms for each subgroup ranged from 11.5% ± 8.0% to 18.6% ± 8.2%. Agreement on the 3 most common signs or symptoms for each subgroup ranged from 2.9% ± 2.1% to 8.6% ± 3.6%. Put simply, at best, only 1 in 10 clinicians could agree on the most common combination of 3 signs or symptoms that indicate these subgroups.

The proportion of clinicians who agreed that the combination of 2 specific signs or symptoms indicated these subgroups was examined by professional discipline. These data show that between 69.2% and 100.0% of the clinicians who agreed on a combination of 2 specific signs and/or symptoms as indicating facet, contained disc, instability, or postural subgroups were either physiotherapists or manipulative physiotherapists. Of the clinicians who agreed that the combination of 2 specific signs or symptoms indicated sacroiliac joint, 50.0% were osteopaths, and 31.3% were chiropractors. Most agreement appeared to be discipline specific.

The results of the cluster analysis (*i.e.*, dendrogram) displayed a smooth transition from solutions with high numbers of clusters continuously down to 2 clusters. These results indicate that cluster analysis of the pool of all the signs and symptoms was unrevealing because heterogeneity in the data prohibited empirical detection of any inherent clusters of signs and symptoms.

## ■ Discussion

### Subgroup Labels

Most clinicians (84%) nominated NSLBP subgroup labels that inferred a putative pathoanatomical source or cause of pain. This pattern was similar across professional disciplines. Almost exclusively, the signs and symptoms that clinicians nominated as being indicative

of subgroups were those derived from the clinical assessment of pain and other physical impairment. The capacity to determine underlying pathoanatomy from these data is controversial. A definitive method to assess whether a cluster of signs and symptoms can correctly identify the source of pain is to compare the results of clinical tests to a gold standard that is known to be a valid test for the source of the pain. There are no gold standards in NSLBP diagnosis, and the accuracy of the available reference standards remains the subject of debate. Presently, provocative diskography is used to investigate disc lesions,<sup>57–60</sup> and anesthetic blocks are used to investigate facet joint pain and sacroiliac joint pain.<sup>60–69</sup> Using these methods, some researchers have concluded that there are no patterns of signs and symptoms that are capable of accurately identifying pathoanatomy in NSLBP.<sup>58,59,66–70</sup> Other researchers have concluded that particular signs and symptoms may have clinical use.<sup>57,60,63,71</sup>

NSLBP pathoanatomical models based on these reference comparisons are controversial as a result of: (1) concerns about the accuracy of inferences regarding primary pain generation because of false-positive rates of provocative diskography and anesthetic injections;<sup>72–74</sup> and (2) concern that workup, selection, and spectrum biases<sup>75</sup> may limit generalizability to a broader clinical population.<sup>76–79</sup>

### Subgroup Composition

There was low agreement regarding the signs or symptoms that are indicative of pathoanatomically named subgroups and a lack of discrete clusters seen in cluster analysis. Therefore, although the specific signs and symptoms nominated for particular subgroups have been illustrated in Table 1 and Appendix 2 (available for viewing on ArticlePlus only), these results should be interpreted with caution.

The low agreement found in the current study is similar to that found in a recent survey of British osteopaths regarding clinical indicators of disc herniation in which



agreement for specific signs and symptoms ranged from 2% to 34%.<sup>80</sup> It is also similar to the agreement (46%) found between rheumatologists diagnosing shoulder pain.<sup>81</sup> In the absence of identifiable pathoanatomy, clinical concordance regarding the signs and symptoms of specific conditions appears to be low.

Most agreement in this study was confined to particular professional disciplines. Reasons for this result might include discipline specific insightful clinical observation, monocultural educational training, and/or an untested discipline specific belief. Furthermore, there were significant differences across disciplines in the reporting rates of particular pathoanatomically named subgroups. This result may reflect the views of different “clinical cultures” or different case mix.

Evidence is not currently available to determine if the views held by any of the professional disciplines are more or less valid than the views held by any other discipline. These data also suggest that most clinicians who responded to this survey use subgroup labels to describe NSLBP heterogeneity and that these labels are recognized across disciplines. It may be that some clinicians believe these labels represent accurate pathoanatomic diagnoses, while other clinicians use these labels as a “shorthand” to communicate a particular cluster of signs and symptoms. However, when using these labels, clinicians may have quite different interpretations as to what that label implies. For example, when one clinician refers to facet syndrome, another clinician’s interpretation of the signs and symptoms associated with this may be quite different. This difference represents a significant barrier to communication within and between the primary care disciplines that treat NSLBP.

The strengths of this survey design and results are: (1) data were gathered from a comprehensive selection of primary care practitioners, and (2) the survey population was comprised of primary care clinicians who deliver the bulk of NSLBP treatment in Australia. The weaknesses of this survey design and results are: (1) the response rate was 59.8%, (2) nonrespondent bias could not be determined, (3) the data relied on the accuracy of clinician self-report, (4) signs and symptoms data were gathered using an open question that required recoding, and (5) the results may not be generalizable to other countries.

Furthermore, in Question 5, the questionnaire asked clinicians to nominate NSLBP subgroups, and the examples given to clarify the intent of the question were “disc lesions, instability, postural syndrome, facet syndrome etc.” These examples were chosen because they were labels commonly nominated by the participants in the pilot study of the questionnaire. The labeling data collected in the questionnaire mimicked the labels derived from the pilot study. This result may be because the pilot study participants nominated labels that are widely used by primary care clinicians. However, it is also possible that the labeling examples provided in Question 5 may have biased the clinician’s responses.

## ■ Conclusions

Most Australian primary care clinicians who responded to this survey think that NSLBP is heterogeneous and treat patients differently based on that heterogeneity. Clinicians think that it is possible to describe the causative pathoanatomy of subgroups within that heterogeneity, and to describe the patterns of signs and symptoms that demarcate those subgroups. However, currently, there is little empirical evidence to support the validity of these views. Furthermore, there is little consensus among clinicians regarding the signs and symptoms that demarcate patterns within that heterogeneity, and the use of descriptive labels, such as facet syndrome, may mask that lack of consensus. NSLBP is a highly prevalent condition for which optimal treatment strategies remain uncertain. There is a clear need for research that examines the outcome validity of subgrouping schemes.

## ■ Key Points

- Of the primary care clinicians who thought that it was currently possible to recognize NSLBP subgroups:
- Nine out of 10 times, clinicians nominated a descriptive label indicating a putative pathoanatomic source of pain.
- There was low agreement regarding the specific signs and symptoms that indicated NSLBP subgroups. At best, only 1 in 10 clinicians agreed on the 3 most common signs and symptoms suggestive of any subgroup.
- Most consensus regarding the specific signs and symptoms that indicated NSLBP subgroups came from within the clinical culture of one professional discipline. There was little consensus across professional disciplines.
- Although the labels given to subgroups were similar across disciplines, descriptive labels such as “Facet Syndrome” should be used with caution because there is little consensus regarding the signs and symptoms indicative of these conditions.

## Acknowledgment

The authors thank the responding clinicians for their participation, and also thank Kathleen Philip, PT, MManipPhysio, and Sue Slade, PT, MManipPhysio, for their independent coding of the data.

Note: Appendices can be viewed online through ArticlePlus only.

## References

1. Deyo R, Rainville J, Kent D. What can the history and physical examination tell us about low back pain. *JAMA* 1992;268:760–5.
2. Spengler DM, David DP. Industrial low back pain: A practical approach. In: Wiesel SW, ed. *Industrial Low Back Pain: A Comprehensive Approach*. Charlottesville, VA: The Michie Co.; 1985:869–71.
3. Kent PM, Keating J. Do primary-care clinicians think that non-specific low back pain is one condition? *Spine* 2004;29:1022–31.

4. Zusman M. Structure-oriented beliefs and disability due to back pain. *Aust J Physiother* 1998;44:13–20.
5. Bogduk N. What's in a name? The labelling of back pain. *Med J Aust* 2000; 173:400–1.
6. Indahl A, Velund L, Reikeraas O. Good prognosis for low back pain when left untampered: A randomized clinical trial. *Spine* 1995;20:473–7.
7. Indahl A, Kaigle A, Reikerås O, et al. Five-year follow-up study of a controlled clinical trial using light mobilization and an informative approach to low back pain. *Spine* 1997;23:2625–30.
8. McGuirk B, King W, Govind J, et al. Safety, efficacy, and cost effectiveness of evidence-based guidelines for the management of acute low back pain in primary care. *Spine* 2001;26:2615–22.
9. Waddell G. *The Back Pain Revolution*. 2nd ed. Edinburgh, NY: Churchill Livingstone; 2004.
10. Bernard T, Kirkaldy-Willis W. Recognizing specific characteristics of non-specific low back pain. *Clin Orthop Relat Res* 1987;217:266–80.
11. Binkley J, Finch E, Hall J, et al. Diagnostic classification of patients with low back pain: Report on a survey of physical therapy experts. *Phys Ther* 1993; 73:138–55.
12. Delitto A, Erhard RE, Bowling RW. A treatment-based classification approach to low back syndrome: Identifying and staging patients for conservative treatment. *Phys Ther* 1995;75:470–89.
13. Fairbank JC, Hall H, van Akkerveken PF, et al. History taking and physical examination: Identification of syndromes of back pain. In: White A, ed. *American Academy of Orthopaedic Surgeons Symposium on Idiopathic Low Back Pain*. St. Louis, MO: CV Mosby; 1982:88–106.
14. Fritz JM. Use of a classification approach to the treatment of 3 patients with low back syndrome. *Phys Ther* 1998;78:766–77.
15. Hall H, McIntosh G, Melles T. A different approach to back pain diagnosis: Identifying a pattern of pain. *Can J CME* 1994:31–41.
16. Maluf K, Sahrmann S, Van Dillen L. Use of a classification system to guide nonsurgical management of a patient with chronic low back pain. *Phys Ther* 2000;80:1097–111.
17. McKenzie R. Mechanical diagnosis and therapy for low back pain: Toward a better understanding. In: Twomey L, Taylor J, eds. *Physical Therapy of the Low Back*. Edinburgh, NY: Churchill Livingstone; 1987:157–73.
18. Moffroid MT, Haugh LD, Henry SM, et al. Distinguishable groups of musculoskeletal low back pain patients and asymptomatic control subjects based on physical measures of the NIOSH Low Back Atlas. *Spine* 1994;19: 1350–8.
19. Mooney V. The classification of low back pain. *Ann Med* 1989;21:321–5.
20. Nelson M. The classification of back syndromes by reliable signs. In: Fairbank JCT, Pynsent PB, eds. *Back Pain: Classification of Syndromes*. Manchester, NY: Manchester University Press; 1990:36–46.
21. Newton W, Curtis P, Witt P, et al. Prevalence of subtypes of low back pain in a defined population. *J Fam Pract* 1997;45:331–6.
22. Saunders H. Classification of musculoskeletal spinal conditions. *J Orthop Sports Phys Ther* 1979;3:3–15.
23. Sikorski JM. A rationalized approach to physiotherapy for low back pain. *Spine* 1985;10:571–9.
24. Spitzer WO, LeBlanc FE, Dupius M. Quebec Task Force on spinal disorders - Scientific approach to the assessment and management of activity-related spinal disorders - a monograph for clinicians. *Spine* 1987;12:S1–S59.
25. Bombardier C, Kerr M, Shannon H, et al. A guide to interpreting epidemiologic studies on the etiology of low back pain. *Spine* 1994;19:20475–56S.
26. Burdorf A, Sorock G. Positive and negative evidence of risk factors for back disorders. *Scand J Work Environ Health* 1997;23:243–56.
27. Frank J, Brooker A, DeMaio S, et al. Disability resulting from occupational low back pain: Part II: What do we know about secondary prevention? A review of the scientific evidence on prevention after disability begins. *Spine* 1996;21:2918–29.
28. Frymoyer J, Cats-Baril W. Predictors of low back pain disability. *Clin Orthop Relat Res* 1987;221:89–98.
29. Hansen DT. Psychosocial predictors in spine care. *Topics Clin Chiro* 1999; 6:38–50.
30. Hogg-Johnson S, Frank J, Rael E. *Prognostic Risk Factor Models for Low-Back Pain: Why They Have Failed and a New Hypothesis*. Toronto, Canada: Institute for Work & Health; 1994:24.
31. Hoogendoorn W, van Poppel M, Bongers P, et al. Systematic review of psychosocial factors at work and private life risk factors for back pain. *Spine* 2000;25:2114–25.
32. Pincus T, Burton A, Vogel S, et al. A systematic review of psychological factors as predictors of chronicity/disability in prospective cohorts of low back pain. *Spine* 2002;27:E109–20.
33. Pincus T, Vlaeyan J, Kendall N, et al. Cognitive-behavior therapy and psychosocial factors in low back pain. *Spine* 2002;27:E133–8.
34. Pulcinis IR, Goel V, Cohen JE, et al. *A Structured Critical Review of Acute Back Pain Prognosis Studies*. Toronto, Ontario, Canada: Institute for Work and Health; 1995:1–26.
35. Teasell RW. Compensation and chronic pain. *Clin J Pain* 2001;17:S46–51.
36. Teasell RW, Bombardier C. Employment-related factors in chronic pain and chronic pain disability. *Clin J Pain* 2001;17:S39–45.
37. Truchon M, Fillion L. Biopsychosocial determinants of chronic disability and low-back pain: A review. *J Occup Rehabil* 2000;10:117–42.
38. Truchon M. Determinants of chronic disability related to low back pain: Towards an integrative biopsychosocial model. *Disabil Rehabil* 2001;23: 758–67.
39. Valat JP, Goupille P, Vedere V. Low back pain: Risk factors for chronicity. *Rev Rhum* 1997;64:189–94.
40. Von Korf M, Deyo RA, Cherkin D, et al. Back pain in primary care. Outcomes at 1 year. *Spine* 1993;18:855–62.
41. Werneke M, Hart DL. Centralization phenomenon as a prognostic factor for chronic low back pain and disability. *Spine* 2001;26:758–65.
42. Werneke M, Hart DL, Cook D. A descriptive study of the centralization phenomenon. A prospective analysis. *Spine* 1999;24:676–83.
43. Long AL. The centralization phenomenon. Its usefulness as a predictor or outcome in conservative treatment of chronic low back pain (a pilot study). *Spine* 1995;20:2513–20.
44. Karas R, McIntosh G, Hall H, et al. The relationship between nonorganic signs and centralization of symptoms in the prediction of return to work for patients with low back pain. *Phys Ther* 1997;77:354–60.
45. Werneke MW, Hart DL. Categorizing patients with occupational low back pain by use of the Quebec task Force Classification System versus pain Classification Procedures: Discriminative and predictive validity. *Phys Ther* 2004;84:243–54.
46. Stankovic R, Johnell O. Conservative treatment of acute low-back pain. A prospective randomized trial: McKenzie method of treatment versus patient education in “mini back school.” *Spine* 1990;15:120–3.
47. Stankovic R, Johnell O. Conservative treatment of acute low back pain. A 5-year follow-up study of two methods of treatment. *Spine* 1995;20:469–72.
48. Fritz JM, Delitto A, Erhard RE. Comparison of classification-based physical therapy with therapy based on clinical practice guidelines for patients with acute low back pain: A randomized clinical trial. *Spine* 2003;28:1363–72.
49. Delitto A, Cibulka MT, Erhard RE, et al. Evidence for use of an extension-mobilization category in acute low back syndrome: A prescriptive validation pilot study. *Phys Ther* 1993;73:216–22.
50. Petersen T, Kryger P, Ekdahl C, et al. The effect of McKenzie Therapy as compared with that of intensive strengthening training for the treatment of patients with subacute or chronic low back pain—A randomized controlled trial. *Spine* 2002;27:1702–9.
51. Cherkin DC, Deyo RA, Battie M, et al. A comparison of physical therapy, chiropractic manipulation, and provision of an educational booklet for the treatment of patients with low back pain. *N Engl J Med* 1998;339:1021–9.
52. Gracey J, McDonough S, Baxter G. Physiotherapy management of low back pain—A survey of current practice in Northern Ireland. *Spine* 2002;27:406–11.
53. Battié MC, Cherkin DC, Dunn R, et al. Managing low back pain: Attitudes and treatment preferences of physical therapists. *Phys Ther* 1994;74:219–26.
54. Foster NE, Thompson KA, Baxter GD, et al. Management of nonspecific low back pain by physiotherapists in Britain and Ireland. A descriptive questionnaire of current clinical practice. *Spine* 1999;24:1332–42.
55. Hair JF, Anderson RE, Ronald LT, et al. *Multivariate Data Analysis*. 5th ed. Upper Saddle River, NJ: Prentice-Hall Intl.; 1998.
56. Tryon WW. Evaluating statistical difference, equivalence, and indeterminacy using inferential confidence intervals: An integrated alternative method of conducting null hypothesis statistical tests. *Psychol Methods* 2001;6:371–86.
57. Donelson R, Aprill C, Medcalf R, et al. A prospective study of centralization of lumbar and referred pain: A predictor of symptomatic discs and annular competence. *Spine* 1997;22:1115–22.
58. Schwarzer A, Aprill C, Derby R, et al. The relative contributions of the disc and zygapophyseal joint in chronic low back pain. *Spine* 1994;19:801–6.
59. Schwarzer AC, Aprill CN, Derby R, et al. The prevalence and clinical features of internal disc disruption in patients with chronic low back pain. *Spine* 1995;20:1878–83.
60. Young S, Aprill C. Characteristics of a mechanical assessment for chronic lumbar facet joint pain. *J Manipulative Physiol Ther* 2000;8:78–84.
61. Fairbank JC, Park WM, McCall IW, et al. Apophyseal injection of local anesthetic as a diagnostic aid in primary low-back pain syndromes. *Spine* 1981;6:598–605.
62. Jackson R, Jacobs R, Montesano P. Facet joint injection in low-back pain—A prospective statistical study. *Spine* 1988;13:966–71.

63. Laslett M, Williams M. The reliability of selected pain provocation tests for sacroiliac joint pathology. *Spine* 1994;19:1243-9.
64. Lilius G, Harilainen A, Laasonen EM, et al. Chronic unilateral low-back pain. Predictors of outcome of facet joint injections. *Spine* 1990;15:780-2.
65. Lippitt A. The facet joint and its role in spine pain—Management with facet joint injections. *Spine* 1984;9:746-50.
66. Revel M, Poiraudau S, Auleley GR, et al. Capacity of the clinical picture to characterize low back pain relieved by facet joint anesthesia. Proposed criteria to identify patients with painful facet joints. *Spine* 1998;23:1972-7.
67. Schwarzer AC, Aprill CN, Derby R, et al. Clinical features of patients with pain stemming from the lumbar zygapophysial joints. Is the lumbar facet syndrome a clinical entity? *Spine* 1994;19:1132-7.
68. Schwarzer AC, Derby R, Aprill CN, et al. Pain from the lumbar zygapophysial joints: A test of two models. *J Spinal Disord* 1994;7:331-6.
69. Schwarzer A, Aprill C, Bogduk N. The sacroiliac joint in chronic low back pain. *Spine* 1995;20:31-7.
70. Manchikanti L, Pampati V, Fellows B, et al. The inability of the clinical picture to characterize pain from the facet joints. *Pain Physician* 2000;3:158-66.
71. Young S, Aprill C, Laslett M. Correlation of clinical examination characteristics with three sources of chronic low back pain. *Spine J* 2003;3:460-5.
72. Carragee EJ, Tanner CM, Khurana S, et al. The rates of false-positive lumbar discography in select patients without low back symptoms. *Spine* 2000;25:1373-80.
73. Carragee EJ, Tanner CM, Yang BB, et al. False-positive findings on lumbar discography: Reliability of subjective concordance assessment during provocative disc injection. *Spine* 1999;24:2542-7.
74. North RB, Kidd DH, Zahurak M, et al. Specificity of diagnostic nerve blocks: A prospective, randomized study of sciatica due to lumbosacral spine disease. *Pain* 1996;65:77-85.
75. Johnston Rea. Assessment of diagnostic tests to inform policy decisions—visual electrodiagnosis. *Int J Technol Assess Health Care* 2003;19:373-83.
76. Begg CB. Biases in the assessment of diagnostic tests. *Stat Med* 1987;6:411-23.
77. Moons KG, van Es GA, Deckers JW, et al. Limitations of sensitivity, specificity, likelihood ratio, and bayes' theorem in assessing diagnostic probabilities: A clinical example. *Epidemiology* 1997;8:12-7.
78. Sackett DL, Haynes RB, Tugwell P. *Clinical Epidemiology*. 1st ed. Boston, MA: Little, Brown and Co.; 1985.
79. Irwig L, Bossuyt P, Glasziou P, et al. Designing studies to ensure that estimates of test accuracy are transferable. *BMJ* 2002;324:669-71.
80. Rebnin R, Baxter D, McDonough S. The passive straight leg raising test in the diagnosis and treatment of lumbar disc herniation: A survey of United Kingdom osteopathic opinion and clinical practice. *Spine* 2003;28:1717-24.
81. Bamji AN, Erhardt CC, Price TR, et al. The painful shoulder: Can consultants agree? *Br J Rheumatol* 1996;35:1172-4.