7.1 Updated Bibliography

A list of all papers referenced in the Clinical Data section, the Executive Summary, the Summary of Safety and Effectiveness, and the Sponsor's Panel Presentation is presented below. Articles shown in boldface indicate that a reprint is provided.

1. Arrington ED, Smith WJ, et al. Complications of iliac crest bone graft harvesting. Clini Orthopaed Rel Res. 1996;329:300-09.

2. Blattert TR, Delling G, Dalal PS, Toth CA, Balling H, Weckbach A. Successful transpedicular lumbar interbody fusion by means of a composite of osteogenic protein-1 (rhBMP-7) and hydroxyapatite carrier: a comparison with autograft and hydroxyapatite in the sheep spine. *Spine.* 2002 Dec 1;27(23):2697-705.

3. **Blumenthal SL, Gill K. Can lumbar spine radiographs accurately determine fusion in postoperative patients? Correlation of routine radiographs with a second surgical look at lumbar fusions. Spine. 1993; 18 (9): 1186-89**.

4. Boden SD, Monroe MA, Martin G, Hutton WC. Posterolateral lumbar spinal arthrodesis with rhBMP-2/HA-TCP following laminectomy in the non-human primate. Trans ORS 1998;19: 249.

5. Boden SD, Schimandle JH, Hutton WC, Chen MI. The use of an osteoinductive growth factor for lumbar spinal fusion. Spine 1995;20(24)2626-2632.

6. Bomback DA, Grauer JN, Lugo R, Troiano N, Patel TCh, Friedlaender GE. Comparison of posterolateral lumbar fusion rates of Grafton Putty and OP-1 Putty in an athymic rat model. *Spine.* 2004 Aug 1;29(15):1612-7.

7. Brandoff JF, Silber JS, Vaccaro AR. Contemporary alternatives to synthetic bone grafts for spine surgery. *Am J Orthop*. 2008 Aug; 37(8):410-4.

8. Bright C, Park YS, Sieber AN, Kostuik JP, Leong KW. In vivo evaluation of plasmid DNA encoding OP-1 protein for spine fusion. *Spine.* 2006 Sep 1;31(19):2163-72.

9. Burkus JK, Gornet MF, Dickman CA, Zdeblick TA. Anterior lumbar interbody fusion using rhBMP-2 with tapered interbody cages. J Spinal Disord Tech. 2002;15(5):337-49.

10. Carlisle E, Fischgrund JS. Bone morphogenetic proteins for spinal fusion. *Spine J.* 2005 Nov-Dec; 5 (6 Suppl):240S-249S.

11. Chen X, Schmidt AH, Tsukayama DT, Bourgeault CA, Lew WD. Recombinant Human Osteogenic Protein-1 Induces Bone Formation in a Chronically Infected, Internally

Stabilized Segmental Defect in the Rat Femur. J Bone Joint Surg Am. (2006) Jul;88(7):1510-23

12. **Carreon LY, Djurasovic M, Glassman SD, Sailer P. Diagnostic accuracy and reliability of fine-cut CT scans with reconstructions to determine the status of an instrumented posterolateral fusion with surgical exploration as reference standard. Spine. 2007 Apr 15;32(8):892-5.**

13. **Carreon LY, Glassman SD, Djurasovic M. Reliability and agreement between fine-cut CT scans and plain radiography in the evaluation of posterolateral fusions. Spine J. 2007 Jan-Feb;7(1):39-43. Epub 2006 Nov 20.**

14. Cook SD, Dalton JE, Tan EH, Whitecloud TS 3rd, Rueger DC. In vivo evaluation of recombinant human osteogenic protein (rhOP-1) implants as a bone graft substitute for spinal fusions. *Spine.* 1994 Aug 1;19(15):1655-63.

15. Cunningham BW, Kanayama M, Parker LM, Weis JC, Sefter JC, Fedder IL, McAfee PC. 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 Mar 15;24(6):509-18.

16. Cunningham BW, Shimamoto N, Sefter JC, Dmitriev AE, Orbegoso CM, McCarthy EF, Fedder IL, McAfee PC. Osseointegration of autograft versus osteogenic protein-1 in posterolateral spinal arthrodesis: emphasis on the comparative mechanisms of bone induction. *Spine J.* 2002 Jan-Feb;2(1):11-24.

17. Einhorn T. Clinical applications of rhBMPs: early experience and future development. *J Bone Joint Surg Am*. 2003;85(A)Suppl 3:82-88.

18. Fehlings et al, J Neurosurg Spine 2007 [See Furlan JC et al., below, for complete reference]

19. **Fischgrund JS, Mackay M, Herkowitz HN, et al. Degenerative lumbar spondylolisthesis with spinal stenosis: a prospective, randomized study comparing decompressive laminectomy and arthrodesis with and without spinal instrumentation. Spine. 1997;22:2807–2812.**

20. 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 Suppl 1(Pt 2):S151-8.

21. **Furlan JC, Perrin RG, Govender PV, Petrenko Y, Massicotte EM, Rampersaud YR, Lewis S, Fehlings MG. Use of osteogenic protein-1 in patients at high risk for spinal pseudarthrosis: a prospective cohort study assessing safety, health-related quality of life, and radiographic fusion. Invited submission from the Joint Section on**

Disorders of the Spine and Peripheral Nerves, March 2007. *J Neurosurg Spine* 2007 Nov; 7(5):486-95.

22. Govender PV, Rampersaud YR, Rickards L, Fehlings MG. Use of osteogenic protein-1 in spinal fusion: literature review and preliminary results in a prospective series of high-risk cases. *Neurosurg Focus.* 2002 Dec 15;13(6):e4.

23. Govender S, Csimma C, Genant HK, Valentin-Opran A. Recominant human bone morphogenetic protein-2 for treatment of open tibial fractures. J Bone Joint Surg. 2002;84-A(12):2123-34.

24. Grauer JN, Vaccaro AR, Kato M, Kwon BK, Beiner JM, Patel TC, Hilibrand AS, Chiba K, Albert TJ. Development of a New Zealand white rabbit model of spinal pseudarthrosis repair and evaluation of the potential role of OP-1 to overcome pseudarthrosis. *Spine.* 2004 Jul 1;29(13):1405-12.

25. Harvinder S. Bone morphogenic proteins and spine surgery. *Spine*. 2003;28(15S):S64-S73.

26. Herkowitz HN, Kurz LT. Degenerative Lumbar Spondylolisthesis with Spinal Stenosis. A Prospective Study Comparing Decompression with Decompression and Intertransverse Process Arthrodesis. J Bone Joint Surg Am 1991 Jul;73(6)802-8.

27. Hisamitsu J, Yamazaki M, Suzuki H, Hashimoto M, Nakajima A, Moriya H. Gene expression for type-specific collagens in osteogenic protein-1 (rhBMP-7)-induced lumbar intertransverse process fusion in rabbits. *Connect Tissue Res.* 2006;47(5):256-63.

28. Imai Y, Okuma M, An HS, Nakagawa K, Yamada M, Muehleman C, Thonar E, Masuda K. Restoration of disc height loss by recombinant human osteogenic protein-1 injection into intervertebral discs undergoing degeneration induced by an intradiscal injection of chondroitinase ABC. *Spine.* 2007 May 15;32(11):1197-205.

29. Jenis LG, Wheeler D, Parazin SJ, Connolly RJ. The effect of osteogenic protein-1 in instrumented and noninstrumented posterolateral fusion in rabbits. *Spine J.* 2002 May-June;2(3):173-8.

30. **Jenis LG, Falb D, Cunningham B, Pierce A, Goad M, Schrier D. Posterolateral Intertransverse Lumbar Fusion in Nonhuman Primates Using OP-1 Putty: Effect of Dose Concentration on Fusion Rate. Proceedings of the NASS 23rd Annual Meeting / The Spine Journal 8 (2008) 1S–191S.**

31. Johnson EE, Urist MR, Finerman GA. Distal metaphyseal tibial nonunion. Deformity and bone loss treated by open reduction, internal fixation, and human bone morphogenetic protein (hBMP). *Clin Orthop Relat Res*. 1990;250:234-40.

32. Johnson EE, Urist MR, Finerman GA. Repair of segmental defects of the tibia with cancellous bone grafts augmented with human bone morphogenetic protein. a preliminary report. *Clin Orthop Relat Res*. 1988;236:249-57.

33. Johnsson R, Strömqvist 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 Dec 1;27(23):2654-61

34. Johnsson KE, Redlund-Johnell I, Uden A, Willner S. Preoperative and postoperative instability in lumbar spinal stenosis. Spine 1989; 14(6):591-593.

35. Kanayama M, Hashimoto T, Shigenobu K, Yamane S, Bauer TW, Togawa D. A prospective randomized study of posterolateral lumbar fusion using osteogenic protein-1 (OP-1) versus local autograft with ceramic bone substitute: emphasis of surgical exploration and histologic assessment. *Spine*. 2006 May 1;31(10):1067-74.

36. Kant AP Daum WJ et al. Evaluation of lumbar spine fusion – plain radiographs versus direct surgical exploration and observation. Spine. 1995; 20(21):2313-17.

37. Kidder LS, Chen X, Schmidt AH, Lew WD. Osteogenic Protein-1 Overcomes Inhibition of Fracture Healing in the Diabetic Rat: A Pilot Study. Clin Orthop Relat Res. 2008 Jul 29. [Epub ahead of print]

38. Kornblum MB, Fischgrund JS, Herkowitz HN, Abraham DA, Berkower DL, Ditkoff JS. Degenerative Lumbar Spondylolisthesis with Spinal Stenosis: A prospective long-term study comparing fusion. Spine 2004; 29(7):726–734

39. Kwiatkowski TC, Meyer RA, Gruber HE, Tabor JR, Murakami T, Howard BA, Wozney JN, Hanley EN. Spinal laminectomy with rhbmp-2 in the canine: a safety study. Trans ORS 1997;18: 189.

40. Lombardi JS, Wiltse LL, Reynolds J, Widell EH, Spencer C. Treatment of degenerative spondylolisthesis. Spine. 1985;10(9) 821-827.

41. Lu J, Bhargav D, Wei AQ, Diwan A. Posterolateral intertransverse spinal fusion possible in osteoporotic rats with BMP-7 in a higher dose delivered on a composite carrier. *Spine*. 2008 Feb 1;33(3):242-9.

42. McAfee PC, Boden SD, et al. Symposium: a critical discrepancy-a criteria of successful arthrodesis following interbody spinal fusions. Spine. 2001;26(3):320-34.

43. MedPAR file 2004. Centers for Medicare and Medicaid Services

44. Mermer MJ, Gupta MC, Wheeler DL, Helgerson J, Reddi AH, Hazelwood S, Benson DR. Efficacy of osteogenic protein-1 in a challenging multilevel fusion model. *Spine.* 2004 Feb 1;29(3):249-56.

45. Millennium Research Group, US Markets for Orthopedic Biomaterials 2005, September 2004.

46. Millennium Research Group, US Markets for Orthopedic Biomaterials 2006, December 2005.

47. Minamide A, Tamakit T, Kawakami M, Hashizume H, Sakatar R. An experimental spinal fusion with wintered bovine bone and recombinant human bone morphogenetic protein-2. Trans ORS 1998;(19): 659.

48. Papakostidis C, Kontakis G, Bhandari M, Giannoudis PV. Efficacy of autologous iliac crest bone graft and bone morphogenetic proteins for posterolateral fusion of lumbar spine: a meta-analysis of the results. *Spine.* 2008 Sep 1;33(19):E680-92.

49. Paramore CG, Lauryssen C, Rauzzino MJ, Wadlington VR, Palmer CA, Brix A, Cartner SC, Hadley MN. The safety of OP-1 for lumbar fusion with decompression-- a canine study. *Neurosurgery.* 1999 May;44(5):1151-5; discussion 1155-6.

50. Patel TC, Erulkar JS, Grauer JN, Troiano NW, Panjabi MM, Friedlaender GE. Osteogenic protein-1 overcomes the inhibitory effect of nicotine on posterolateral lumbar fusion. *Spine.* 2001 Aug 1;26(15):1656-61.

51. Poynton AR, Lane JM. Safety profile for the clinical use of bone morphogenetic proteins in the spine. *Spine.* 2002 Aug 15;27(16 Suppl 1):S40-8.

52. Russell JL, Block JE. Surgical harvesting of bone graft from the ilium: point of view. Medical Hypotheses. 2000;55(6):474-479.

53. Sandhu HS, Khan SN. Animal models for preclinical assessment of bone morphogenetic proteins in the spine. Spine. 2002;15;27(16 Suppl 1):S32-8. Review.

54. Sandhu HS, Kanim LE, Kabo JM, Toth JM, Zeegen EN, Liu D, Delamarter RB, Dawson EG. Effective doses of recombinant human bone morphogenetic protein-2 in experimental spinal fusion. Spine. 1996 Sep 15;21(18):2115-22.

55. Seiler JG III, Johnson J. Iliac crest autogenous bone grafting: donor site complications. J South Orthop Assoc. 2000;9(2):91-97.

56. Singh, K, Smucker, J., Boden, S. Use of Recombinant Human Bone Morhogenetic Protein-2 as an Adjunct in Posterolateral Lumbar Spine Fusion: A Prospective CT-Scan Analysis at One and Two years. J Spinal Disord Tech 2006; 19(6):416-423

57. Summers BN, Eisenstein SM. Donor site pain from the ilium - a complication of lumbar spine fusion. J Bone Joint Surg Am. 1989;71-B(4).

58. Tsutsumimoto T, Shimogata M, Yoshimura Y, Misawa H. Union versus nonunion after posterolateral lumbar fusion: a comparison of long-term surgical outcomes in patients with degenerative lumbar spondylolisthesis. *Eur Spine J*. 2008 Aug;17(8):1107-12. Epub 2008 Jun 7.

59. Urist MR. Bone: formation by autoinduction. *Science*. 1965;150:893-9.

60. Vaccaro AR, Patel T, Fischgrund J, Anderson DG, Truumees E, Herkowitz H, Phillips F, Hilibrand A, Albert TJ. A 2-year follow-up pilot study evaluating the safety and efficacy of op-1 putty (rhbmp-7) as an adjunct to iliac crest autograft in posterolateral lumbar fusions. *Eur Spine J.* 2005 Sep;14(7):623-9. Epub 2005 Jan 26.

61. Vaccaro AR, Patel T, Fischgrund J, Anderson DG, Truumees E, Herkowitz HN, Phillips F, Hilibrand A, Albert TJ, Wetzel T, McCulloch JA. A pilot study evaluating the safety and efficacy of OP-1 Putty (rhBMP-7) as a replacement for iliac crest autograft in posterolateral lumbar arthrodesis for degenerative spondylolisthesis. *Spine*. 2004 Sep 1;29(17):1885-92.

62. Vaccaro AR, Anderson DG, Toth CA. Recombinant human osteogenic protein-1 (bone morphogenetic protein-7) as an osteoinductive agent in spinal fusion. *Spine.* 2002 Aug 15;27(16 Suppl 1):S59-65.

63. Vaccaro AR, Anderson DG, Patel T, Fischgrund J, Truumees E, Herkowitz HN, Phillips F, Hilibrand A, Albert TJ, Wetzel T, McCulloch JA. Comparison of OP-1 Putty (rhBMP-7) to iliac crest autograft for posterolateral lumbar arthrodesis: a minimum 2-year follow-up pilot study. *Spine*. 2005 Dec 15;30(24):2709-16.

64. Vaccaro AR, Patel T, Fischgrund J, Anderson DG, Truumees E, Herkowitz H, Phillips F, Hilibrand A, Albert TJ. 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 Oct;12(5):495-500. Epub 2003 Aug 8.

65. **Vaccaro AR, Lawrence JJP, Patel T, Katz LD, Anderson DG, Fischgrund JS, Krop J, Fehlings MG, Wong D. The Safety and Efficacy of OP-1 (rhBMP-7) as a Replacement for Iliac Crest Autograft in Posterolateral Lumbar Arthrodesis: A Long Term (>4 years) Pivotal Study. *Spine.* 2008 33(26): 2850–2862**.

66. Vaccaro AR, Chiba K, Heller JG, Patel TCh, Thalgott JS, Truumees E, Fischgrund JS, Craig MR, Berta SC, Wang JC; North American Spine Society for Contemporary Concepts in Spine Care. Bone grafting alternatives in spinal surgery. Spine J. 2002 May-Jun;2(3):206-15. Review.

67. Valentin-Opran A, Wozney J, Csimma C, Lilly L, Riedel, GE. Clinical evaluation of rhBMP-2. Clin Orthop Relat Res. 2002;3:79-83.

68. Weinsten, J, et al. The Impact of Comorbidities on the Change in Short-Term 36 and Oswestry Scores Following Lumbar Spine Surgery. Spine 2006; 31:1974-1980.

69. White AP, Maak TG, Prince D, Vaccaro AR, Albert TJ, Hilibrand AS, Grauer JN. Osteogenic protein-1 induced gene expression: evaluation in a posterolateral spinal pseudarthrosis model. *Spine.* 2006 Oct 15;31(22):2550-5.

70. Yao G, Qian Y, Chen J, Fan Y, Stoffel K, Yao F, Xu J, Zheng MH. Evaluation of insoluble bone gelatin as a carrier for enhancement of osteogenic protein-1-induced intertransverse process lumbar fusion in a rabbit model. *Spine.* 2008 Aug 15;33(18):1935-42.

71. Resnick K, Choudhri Dailey A, et al. Guidelines for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 4: radiographic assessment of fusion. J Neurosurg: Spine 2005; 2:653–657.

SPINE Volume 18, Number 9, pp 1186–1189
©1993, J. B. Lippincott Company

Can Lumbar Spine Radiographs Accurately Determine Fusion in Postoperative Patients?

Correlation of Routine Radiographs with a Second Surgical Look at Lumbar Fusions

Scott L. Blumenthal, MD, and Kevin Gill, MD

Controversy centers on the determination of surgical fusion in lumbar spinal patients. What method best determines the nature of surgical arthrodesis remains unanswered. Numerous studies have investigated the accuracy of different radiologic tests. Although the best method has not been determined, plain radiography is certainly the most widely used in many centers and reported in scientific articles. In most of the literature a poor agreement between radiographic interpretation and surgical findings was observed. The main reasons seemed to be the lack of an accurate method to assess the radiographs. The authors rely on a radiologist or spinal surgeon to estimate the success of bony fusion by reviewing a two-dimensional radiograph. The purpose of this study was to determine the accuracy of plain radiographs to predict the presence of a surgical fusion. Forty-nine patients underwent fusion site exploration in the course of hardware removal. All patients had a one- or two-level posterolateral fusion and posterior lumbar interbody fusion with pedicle screw/link rod instrumentation. Immediate preoperative anteroposterior and lateral radiographs were taken before hardware removal that included both visual assessment and a kocher mechanical test. Two spinal surgeons and two musculoskeletal radiologists blindly judged the preoperative radiographs as to the absolute presence or absence of successful arthrodesis. A second review was repeated at 3 months. The overall agreement between radiographic assessment and actual surgical findings was 69%. The range among observers was 57–77%. The overall false positive rate was 42% (0–75%), while the false-negative rate was 29% (20–51%). Success of observed surgical arthrodesis at the time of the second look was 90% and this number was used as the standard in the agreement process. In comparing the radiographic observations with the surgical findings it is suggested that in one of five cases the plain radiographs underestimate the degree of fusion. This finding agrees with the authors' knowledge of osteoid and mineralized bone. The premineralized osteoid may be functionally fused, but appear radiolucent on radiographic film. Once solid trabecular bony bridging occurs radiographic identification of fusion is easier to determine. [Key words: fusion, radiographic studies, lumbar spine]

The goal of spinal fusion is quite obvious: to obtain a solid arthrodesis. Difficulties in determining that outcome are emphasized in a recent publication by Brodsky.[2] They maintain that "Bosworth's Dictum (the only way to be sure of the status of fusion was to explore it) still holds."[1,3] Brodsky's report correlates preoperative studies including plain radiographs, computed tomography, tomography, and bending radiographs with surgical exploration of lumbar spine fusions. Overall plain radiographs exhibited the greatest accuracy in prediction of fusion ranging from 64 to 77%.

Although it is not universally agreed that plain radiographs are the best predictors of arthrodesis they are certainly the most widely used. In fact, many scientific articles reporting outcomes of spinal fusion have used this modality to report their fusion rate. Radiographs relying on two-dimensional representation of three-dimensional anatomy can be problematic at best for this purpose.

We undertook this study to correlate the postoperative radiologic studies to the intraoperative finding of the presence or absence of arthrodesis.

Materials and Methods

The population studied included 49 patients (17 women and 32 men) with an age range of 22–54 years, who underwent one- or two-level lumbar interbody fusions combined with posterolateral fusions between 1988 and 1990. All patients were stabilized with the Long Beach System® (Wiltse System). At an average of 9 months after surgery, elective hardware removal was undertaken. The primary indication for hardware removal was persistent lumbosacral discom-

From the Southwest Orthopedic Institute, Dallas, Texas.
Accepted for publication July 28, 1992.

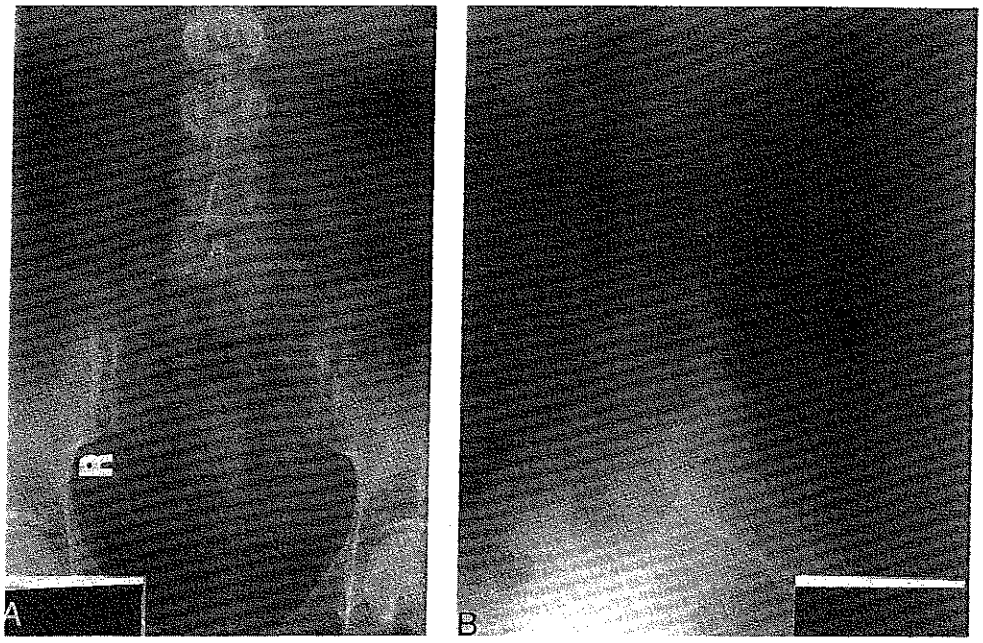


Figure 1. **A**, 38-year-old woman after posterior lumbar interbody fusion (PLIF) and posterolateral fusion (PLF) of L5/S1. This anteroposterior radiograph shows solid PLF that was confirmed at surgery. **B**, Lateral radiograph illustrating solid interbody fusion at L5/S1.

fort. In some cases the patient's wish to have the instrumentation removed was fear of metallic presence in their body. An intraoperative assessment of arthrodesis was made by two methods. The bony mass was explored bilaterally to determine the presence or absence of a bridging bony mantle. In addition, a mechanical stress test of the motion segment was performed with a Kocher clamp and observation of any motion was made and recorded.

Immediate preoperative anteroposterior and lateral radiographs of these patients were gathered and encoded to blind the observers to the identities of the patients. Fusion assessment was then made by two spinal surgeons and two experienced musculoskeletal radiologists. Figures 1–4 illustrate examples of each possible combination of observed radiographs and fusions. Figure 1 illustrates an example in which all observers noted that fusion appeared solid on radiograph and surgery confirmed fusion. Figure 2 illustrates an example in which none of the observers believed that fusion was present and pseudarthrosis was confirmed at surgery. Figure 3 illustrates an example in which all reviewers thought pseudarthrosis was present, yet surgery confirmed solid arthrodesis (false-negative). Figure 4 illustrates an example in which three of four observers thought that fusion was present and pseudarthrosis was confirmed at surgery (false-positive). A second radiographic review was repeated by all four observers at 3 months. Data were then recorded, correlated, and evaluated by a statistician.

Results

Results are presented based on correlations of the radiographic findings and surgical assessment of fusion (Table 1) Solid fusion is considered a positive finding and pseudarthrosis a negative finding. Fusion was observed in 44 of 49 patients or 90%. This was used as

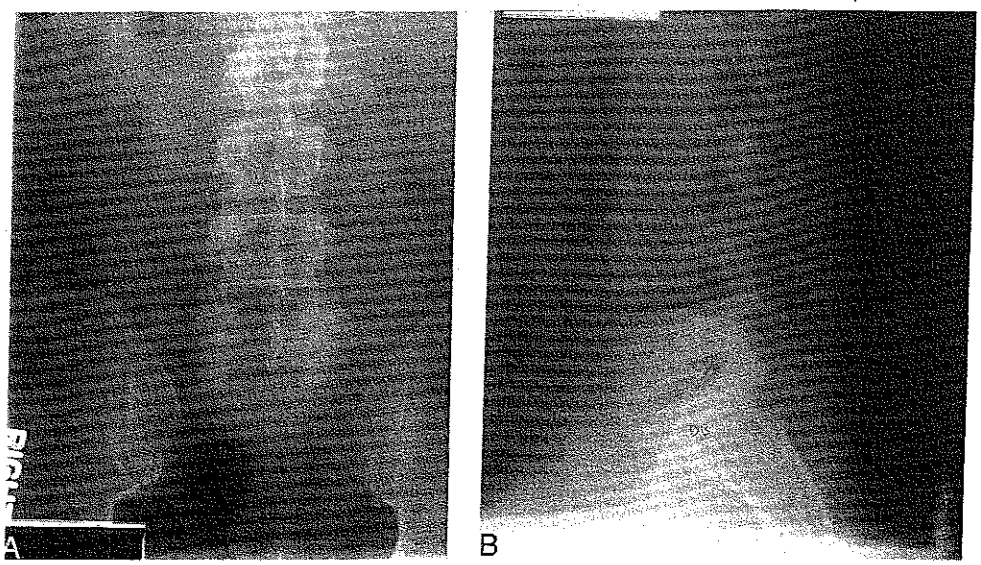


Figure 2. **A**, 40-year-old man after posterior lumbar interbody fusion and posterolateral fusion at L5/S1. This anteroposterior radiograph illustrates failure of posterolateral fusion. **B**, Lateral radiograph demonstrating pseudarthrosis of the interbody bone graft L5/S1.

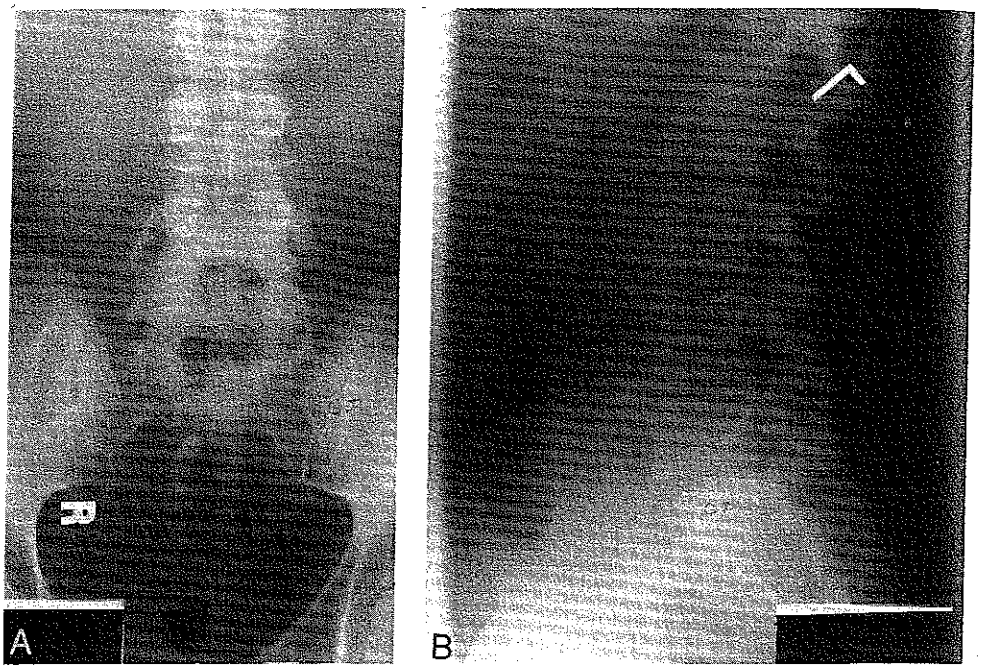


Figure 3. A 35-year-old man after anterior lumbar interbody fusion and posterolateral fusion at L4/5 and L5/S1, respectively. This anteroposterior radiograph appears to illustrate pseudarthrosis of the posterolateral fusion mass, however, surgery confirmed solid fusion mass in the posterolateral area. **B**, Lateral radiograph demonstrating possible pseudarthrosis of the anterior lumbar interbody fusion L4/5 and L5/S1.

the gold standard in the agreement process of statistics. On both the first and second ratings, this is an obvious range of inter-rater consistencies for both the first (67%–77%) and the second (57%–73%) ratings. Moreover, there is a great degree of variability in sensitivity and specificity rates among the raters. One neuroradiologist had a false-positive rate of zero suggesting a more critical definition of bony arthrodesis.

The range of intra-rater consistencies in assessing the radiographs on two separate occasions is seen in Table 2. Cohen's Kappa is a chance corrected estimate of agreement and reflects how much better than chance the intra-rater consistency is. There is an obvious range of intra-rater consistencies (71%–90%), however, all are within the fair to good range of agreement beyond chance.

Table 1. Range of Consistencies Between Ratings of Radiographs and Actual Surgical Findings

	Raters			
	1	2	3	4
First rating				
% consistency between rating and actual surgical finding	71	77	69	67
False-positive rate (%)	38	38	38	75
False-negative rate (%)	27	20	29	24
Sensitivity	0.73	0.81	0.71	0.76
Specificity	0.63	0.63	0.63	0.25
Second rating				
% consistency between rating and actual surgical finding	67	73	57	69
False-positive rate (%)	50	25	0	75
False-negative rate (%)	29	27	51	22
Sensitivity	0.71	0.73	0.49	0.78
Specificity	0.50	0.75	1.00	0.25

Sensitivity refers to the probability of raters judging "positive" when the case or criterion is truly positive, whereas specificity refers to the probability of raters judging "negative" when the case or criterion is truly negative.

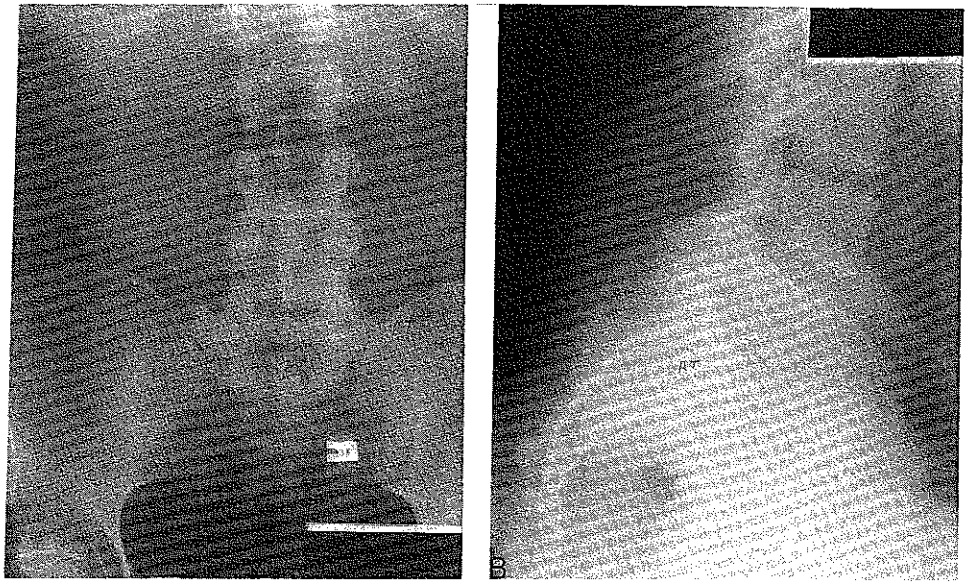


Figure 4. **A**, 40-year-old man after posterior lumbar interbody fusion and posterolateral fusion at L5/S1. This anteroposterior radiograph illustrates possible fusion of the posterior fusion mass recorded by three of four observers, however surgery confirmed pseudarthrosis. **B**, Lateral radiograph illustrating possible fusion of allograft posterior lumbar interbody fusion at L5/S1.

Discussion

In our series, the overall agreement between radiographic assessment of fusion and actual surgical results was 69%. This number best correlates with group I in Brodsky's series.[2] Their plain radiographic correlation rate was 64%, the highest of all modalities used for fusion assessment.

It is recognized that plain radiographs judge the structural integrity and do not realize the functional integrity as do flexion–extension radiographs. Intraoperatively both structural (direct observation) and functional integrity (motion assessment) were recorded. Because all of the patients in this series were stabilized with rigid internal fixation it is unlikely that preoperative flexion–extension radiographs would be of significant value in terms of judging the functional integrity of the fusion mass. However, it is yet to be determined whether the Wiltse system with its fixed link-rod assembly will permit limited motion capable of assessment with flexion–extension radiographs.

In painful lumbar degenerative disc disease, determination of the fusion status becomes important only in patients with increasing deformity or symptoms. The results of this and previous studies failed to indicate a high level of accuracy in the radiographic prediction of fusion. We agree with Brodsky that routine exploration of fusion is not recommended. However, in cases of persistent or increasing symptoms, exploration to assess and perhaps treat problems with the fusion construct may be indicated.

What can be gleaned from this study is that when using static two-dimensional radiographs one can predict the presence or absence of arthrodesis in roughly two thirds of cases. Furthermore, in one of five cases the degree of fusion was underestimated. It is recognized with increasing time that the process of revascularization allows further mineralization of a fusion and this may enhance the radiographic appearance of a solid bony mass. Relative to this, both structural as well as functional aspects of fusion become an important consideration. More importantly this must be correlated with the patient's clinical course to determine if further treatment is advisable.

Table 2. Range of Intrarater Consistencies in Rating of Radiographs on Two Separate Occasions: Three-Month Interval

	Raters			
	1	2	3	4
% Intrarater consistency	76	80	71	90
Cohen's kappa	0.44	0.62	0.42	0.72

References

1. Bosworth DM: Techniques of spinal fusion: Pseudarthrosis and method of repair. AAOS Instructional Course Lecture V. 1948, pp 295–313
2. Brodsky AE, Evan SK, Momtaz AK: Correlation of radiologic assessment of lumbar spine fusions with surgical exploration. Spine 16:S261–265, 1991
3. Cleveland M, Bosworth DM, Thompson FR: Pseudarthrosis in the lumbosacral spine. J Bone Joint Surg 30A:302–312, 1948

Address reprint requests

Kevin Gill, MD
5920 Forest Park Rd, LB 560
Dallas, TX 75235

SPINE Volume 32, Number 8, pp 892–895
©2007, Lippincott Williams & Wilkins, Inc.

Diagnostic Accuracy and Reliability of Fine-Cut CT Scans With Reconstructions to Determine the Status of an Instrumented Posterolateral Fusion With Surgical Exploration as Reference Standard

Leah Y. Carreon, MD, MSc,* Mladen Djurasovic, MD,*† Steven D. Glassman, MD,*† and Philip Sailer, MD*†

Study Design. Accuracy of a diagnostic test referenced to the gold standard.

Objectives. This study evaluated the reliability and accuracy of fine-cut computed tomography scans with coronal and sagittal reconstructions to determine the status of an instrumented posterolateral fusion by using surgical exploration as the reference standard.

Summary of Background Data. There is still a need for a reliable and accurate noninvasive method to determine the status of a spinal fusion.

Methods. Three spine surgeons reviewed 93 prerevision fine-cut CT scans over 163 fused levels of consecutive patients who had revision surgery after an instrumented posterolateral lumbar fusion. The facet joints and posterolateral gutters at each level were classified as fused or not. The surgeons were unaware of the findings on surgical exploration. Interobserver variability and likelihood ratios for a solid fusion when both, one, or none of the facets and when both, one, or none of the posterolateral gutters were fused, were calculated.

Results. There were 42 males and 51 females with a mean age of 57 years (range, 19–86 years) at revision. On exploration, there were 32 (19.6%) nonunions over 163 levels. The kappa for interobserver variability for evaluating facet fusions (0.42) was moderate and for posterolateral fusions (0.62) was substantial. The probability of a solid fusion on exploration was higher when both posterolateral gutters were fused on CT scan (89%) than when both facets were fused on CT scan (74%). When both facets and both posterolateral gutters were fused on CT scan, the probability of a solid fusion on exploration is 96%. The absence of fusion of one or both facets or one posterolateral gutter were poor predictors of nonunion on surgical exploration.

Conclusions. The CT scan reading of either one or both posterolateral gutters fused or both facets fused were moderately predictive of a solid fusion on surgical exploration. Fine-cut CT scans with reconstructions are moderately predictive of the presence of nonunion when both facets are not fused.

From the *Kenton D. Leatherman Spine Center, Louisville, KY; and †Department of Orthopaedic Surgery, University of Louisville School of Medicine, Louisville, KY.
Acknowledgment date: February 21, 2006. First revision date: April 14, 2006. Second revision date: July 7, 2006. Acceptance date: August 31, 2006.
Supported by Norton Healthcare.
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 reprint requests to Leah Y. Carreon, MD, MSc, Kenton D. Leatherman Spine Center, 210 East Gray Street, Suite 900, Louisville, KY 40202; E-mail: lcarreon@spinemds.com

Key words: computed tomography, diagnostic accuracy, posterolateral fusion, likelihood ratio. **Spine 2007; 32:892–895**

Surgical exploration continues to be the gold standard to evaluate the status of a posterolateral fusion. The need for a reliable and accurate noninvasive method to determine the status of a spinal fusion is underscored by recent controversy on the efficacy of spinal fusion in the treatment of lumbar degenerative conditions. Clinically, in patients who have undergone a lumbar fusion and have new or recurrent symptoms, determining the presence of a solid fusion also becomes important.

CT scan, with its ability to provide osseous detail, has become the diagnostic imaging of choice to evaluate spine fusions. Previous studies on the accuracy of CT scans using 6-mm axial slices reported a 57% correlation with surgical exploration.[1] The addition of selective sagittal reconstructions increased this correlation to 80%.[2]

Current high-resolution CT scanners produce contiguous 1-mm-thick axial sections with a 1-mm table increment to optimize spatial resolution and to enhance the quality of computer-generated reformatted images. Contiguous axial images and contiguous reformatted sagittal and coronal images decrease the probability that bony bridging over a very limited area will be missed. Thinner sections also provide improved spatial resolution and improve the quality of reformatted images.

This study was conducted to determine the reliability and accuracy of fine-cut CT scans with coronal and sagittal reconstructions to determine the status of an instrumented posterolateral fusion by using surgical exploration as the reference standard. The status of the both facets and posterolateral gutters on either side were evaluated.

Methods

Three spine surgeons retrospectively reviewed prerevision fine-cut CT scans of consecutive patients who had revision surgery after an instrumented posterolateral lumbar fusion from April 2002 to August 2005. Patients were seen at a multisurgeon spine surgery specialty clinic. Indications for revision included nonunion, painful instrumentation and adjacent level degeneration. A total of 109 patients met the inclusion criteria. Sixteen studies were excluded due to missing films, leaving 93 CT scans with 163 fused levels for evaluation.

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

The right and left facet joint and the right and left posterolateral gutters at each level were classified as fused or not. A facet fusion was defined as obliteration of the joint space between the superior and inferior articulating surfaces of the facet. A posterolateral gutter fusion was defined as continuous trabeculated bone connecting the transverse processes. If the fusion was doubtful or probable, it could not be classified as fused. The surgeons reviewing the CT scans were not aware of the findings on surgical exploration and were not involved in the care of the patients.

All CT scans were performed at the same institution using the same technique. CT scans were 1 mm thick, continuous, nonoverlapping axial slices. The gantry was tilted to obtain scans parallel to the disc space and stayed constant throughout the scan. The field of view included all fused vertebrae to include all transverse processes. Window and level settings were 2000/350 on the GE scanners (General Electric) to optimize trabecular bone detail. All the films reviewed were hard copies and not computer-assisted filmless digital images.

Surgical exploration was done through an open posterior midline approach. The screws were uncoupled from the rods and the rods removed. Distraction forces were then applied over the screw heads using a laminar spreader to detect any motion across the fusion mass. After removal of the screws, pedicle probes were inserted into the pedicles and distraction force was again applied using a laminar spreader to detect any motion across the fusion mass. Inspection of the fusion masses was done to establish the presence or absence of bony continuity. Absence of bony continuity on inspection of the posterolateral gutters and facets and the presence of motion on distraction across the fused levels were defined as a nonunion on surgical exploration.

Interobserver variability in the evaluation of the CT scans for fusion across the facets and the posterolateral gutters was computed using kappa coefficients. The strength of the agreement was interpreted based on the classification of Landis and Koch.[3] The likelihood ratios for a solid fusion when both, one, or none of the facets and when both, one or none of the posterolateral gutters were fused were calculated for each of the 3 raters as well as the group as a whole. The likelihood ratios for a nonunion when both, one, or none of the facets and when both, one, or none of the posterolateral gutters were fused were also computed. The rating used to calculate likelihood ratios for the group was the consensus opinion of the 3 raters. If at least 2 raters thought that the patient was fused, then the group rating was defined as fused.

Likelihood ratios were calculated instead of sensitivity and specificity since there are 3 possible outcomes for each parameter, facet fusion, and posterolateral fusion. The likelihood ratio is a ratio of 2 probabilities, the probability of a given test result among people with a disease divided by the probability of that test result among people without the disease. In this study, the likelihood ratio for a solid fusion is the probability of a given CT scan reading (both, one, or none of the facets or posterolateral gutters were fused) among patients with a solid fusion on exploration divided by the probability of the test result among patients with a nonunion on exploration. The likelihood ratio for a nonunion is the probability of a given CT scan reading (both, one, or none of the facets or posterolateral gutters were fused) among patients with a nonunion on exploration divided by the probability of the test result among patients with a solid fusion on exploration. Likelihood ratios are interpreted as follows: likelihood ratios greater than 10 or less than 0.1 generate large changes in probabilities, ratios from 5 to 10 and 0.1 to 0.2 generate moderate changes, ratios from 2 to 5 and 0.5 to 0.2 generate small changes, and ratios from 1 to 2 and 0.5 to 1 generate extremely small changes.[4]

Table 1. Preoperative Indications for Surgery and Postoperative Diagnosis in Each Patient

	Postoperative Diagnosis			
Preoperative Indication for Surgery	Nonunion	Adjacent Level Degeneration	Painful Instrumentation	Total
Nonunion	25	7	2	34
Adjacent level degeneration	5	47	1	53
Painful instrumentation	2	3	1	6
Total	32	57	4	93

■ Results

There were 42 males and 51 females with a mean age of 57 years (range, 19–86 years) at the time of revision surgery. There were 38 smokers. Thirty-two (34%) required revision surgery for repair of a nonunion, 57 (61%) had adjacent level degeneration, and 4 (4%) had painful instrumentation. Of the 32 patients who had nonunion repair, only 26 had a preoperative diagnosis of nonunion. The preoperative diagnosis was adjacent level degeneration in 4 and painful instrumentation in 2 (Table 1). Seventeen of the 32 nonunions were smokers. Forty-two patients had single-level fusions, 37 had 2-level, 13 3-level, and 2 4-level fusions. There was an average of 49 months (±38 months) between the initial and the revision surgery (range, 1–148 months) and an average of 4 months (±3 months) between the date the CT scan was taken and the revision surgery (range, 1–23 months). There were 32 nonunions over 163 levels on exploration, giving a prevalence of 19.6%.

The kappa coefficient for interobserver variability for evaluating facet fusions (0.42) showed moderate agreement between the raters, while the agreement for posterolateral gutter fusion was substantial (kappa = 0.62). The results of the group consensus for the status of the facet fusion are presented in Table 2, and the results of group consensus for the status of the posterolateral gutter fusion are presented in Table 3.

The likelihood ratios for a solid fusion for each of the raters and the group are presented in Table 4. When both facets were read as fused, it was 2.90 times more likely

Table 2. Group Consensus for the Status of the Facet Fusion

	Findings on Surgical Exploration	
CT Scan Evaluation	Fused	Not Fused
Both facets fused	107	9
One facet fused	9	4
No facet fused	15	19
Total	131	32

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

Table 3. Group Consensus for the Status of the Posterolateral Gutter Fusion

	Findings on Surgical Exploration	
CT Scan Evaluation	Fused	Not Fused
Both gutters fused	68	2
One gutter fused	22	1
No gutter fused	41	29
Total	131	32

Table 5. Positive Predictive Values for a Solid Fusion

	Predictive Value for Solid Fusion			
CT Scan Evaluation	Rater 1 (%)	Rater 2 (%)	Rater 3 (%)	Consensus (%)
Both facets fused	77	71	73	74
One facet fused	47	55	39	35
No facet fused	12	29	13	16
Both gutters fused	85	85	89	89
One gutter fused	84	56	49	84
No gutter fused	27	26	28	26

that the patient was fused on surgical exploration, when only one facet was read as fused it was 0.55; and when both facets were not fused, it was 0.19. When both posterolateral gutters were read as fused, it was 8.31 times more likely that the patient was fused on surgical exploration; when one posterolateral gutter was not fused, it was 5.37; and when both posterolateral gutters were not fused, it was 0.35.

Given an 80.4% prevalence of solid fusion per level in this study sample, the probability of a patient having a solid fusion on surgical exploration was higher (Table 5) when one (84%) or both (89%) posterolateral gutters were fused on CT scan than when both facets were fused (74%).

The likelihood ratios for nonunion for each of the raters and the group are presented in Table 6. When neither facet was fused, it was 5.19 times more likely that a nonunion was present on surgical exploration, when one facet was fused it was 1.82, and when both facets were fused it was 0.34. When neither posterolateral gutter was fused, it was 2.90 times more likely that the patient had a nonunion on surgical exploration; when one posterolateral gutter was fused it was 0.19, and when both posterolateral gutters were fused it was 0.12. With a per level nonunion prevalence of 19.6%, a CT scan finding of the absence of fusion of both facets gave a higher probability of nonunion on surgical exploration (84%) than the absence posterolateral gutter fusion (Table 7).

■ Discussion

Determination of the fusion status in patients who have had a lumbar fusion presenting new or recurrent symptoms is important. Previous studies evaluated plain radiographs,[5–7] bending films,[1,6] stereophotogrammetry,[8] radiograph polytomography,[1,9] magnetic resonance imaging, and radionuclide imaging[6,9] as diagnostic tools to predict the status of a lumbar fusion. Various techniques of computed axial tomography,[1,2,6] including 2-dimensional multiplanar reformations or 3-dimensional reconstructions[10,11] and direct coronal CT scanning,[12] have been used to assess fusion status with varying success. Larsen *et al*[6] reported that CT scans predicted the surgical result in 15 of 24 patients. Brodsky *et al*[1] reported on 175 patients with posterolateral spinal fusions with and without spinal instrumentation, who subsequently underwent surgical assessment of their fusions. They showed a 57% correlation between fusion assessment using 6-mm axial slice CT scans and surgical exploration. This low correlation may be explained by the thickness of the axial slices as well as the lack of 2-dimensional multiplanar reformations in the sagittal and coronal planes. Laasonen and Soini,[2] using 6-mm CT scan with selective sagittal reconstructions in 20 patients found an 80% correlation between findings on the CT scan and surgical exploration. In both these studies, the CT scans were not done specifically to assess the status of the spinal fusion and the thickness of the axial sections could have limited the resolution and diagnostic information available in these imaging studies, even when reformatting was performed. Our current study evaluated CT scans with contiguous 1-mm-thick axial sections with a 1-mm table increment and reformatted sagittal and coronal images were obtained. These CT scans were taken specifically to assess the fusion in patients who were symptomatic.

Evaluating the facets and the posterolateral gutters on each side separately allowed the raters to focus on these 4 areas where the fusion procedure is surgically per-

Table 4. Likelihood Ratios for a Solid Fusion

	Likelihood Ratio for Solid Fusion			
CT Scan Evaluation	Rater 1	Rater 2	Rater 3	Consensus
Both facets fused	3.30	2.42	2.74	2.90
One facet fused	0.90	1.22	0.64	0.55
No facet fused	0.14	0.41	0.15	0.19
Both gutters fused	5.46	5.78	8.31	8.31
One gutter fused	5.13	1.27	0.94	5.37
No gutter fused	0.38	0.35	0.38	0.35

Table 6. Likelihood Ratios for Nonunion

	Likelihood Ratio for Nonunion			
CT Scan Evaluation	Rater 1	Rater 2	Rater 3	Consensus
No facet fused	7.16	2.43	6.82	5.19
One facet fused	1.12	0.82	1.56	1.82
Both facets fused	0.30	0.41	0.36	0.34
No gutter fused	6.38	2.89	2.62	2.90
One gutter fused	0.19	0.79	1.06	0.19
Both gutters fused	0.18	0.04	0.12	0.12

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

Table 7. Positive Predictive Values for Nonunion

	Predictive Value for Nonunion			
CT Scan Evaluation	Rater 1 (%)	Rater 2 (%)	Rater 3 (%)	Consensus (%)
No facet fused	88	71	87	84
One facet fused	53	45	61	65
Both facets fused	23	29	27	26
No gutter fused	86	74	72	74
One gutter fused	16	44	51	16
Both gutters fused	15	4	11	11

formed and assess the adequacy of the fusion. Even though the CT scans were performed in a single institution using the same imaging protocol, and all the raters were fellowship trained spine surgeons and were given specific instructions on how to rate the fusion, the interobserver reliability for evaluating facet fusions was only moderate. This may be due to the variability in the amount of facet resected during decompression and the orientation of the facet joint in relation to the frontal and sagittal plane CT reformations. There was substantial agreement among the observers on the evaluation of the posterolateral gutters, which may be reflective of the ease in which bony continuity can be assessed with coronal and sagittal plane reformations.

The likelihood ratio is a measure of diagnostic accuracy that indicates how much the result of a diagnostic test, in this case the CT scan findings, increase or decrease the probability of the disease. In this study, the likelihood ratios presented in Table 4 show what the odds are that the patient has a solid fusion on surgical exploration, given the different CT scan findings. The CT scan reading of either one or both posterolateral gutters fused were moderately predictive of a solid fusion on surgical exploration. CT scan findings based on the status of the facet fusion were not predictive of the presence or absence of a solid fusion on exploration.

Table 6 shows the likelihood ratios for nonunion for the different CT scan results using the group consensus. The likelihood ratio for no facet fusion is moderately predictive of the presence of nonunion on surgical exploration. The likelihood ratio for both posterolateral gutter fused is moderately predictive of the absence of a nonunion.

This study has several limitations. Foremost is that only symptomatic patients who required a revision surgery for nonunion, adjacent level degeneration, and painful instrumentation were included. The ideal study design would be to perform fine-cut CT scans on all patients who had a posterolateral fusion, symptomatic or not, followed by a surgical exploration. However, this is not ethically possible. Another limitation is that the instrumentation used in all patients was titanium. Thus, the results of this study may not be replicable in cases were stainless steel screws were used, which create artifacts on CT scans.

Conclusion

Fine-cut CT scans with coronal and sagittal reconstructions can reliably and can predict the presence of a solid posterolateral fusion with moderate accuracy if either one or both posterolateral gutters show continuous bony trabeculations or if both facet joints are obliterated. Fine-cut CT scans with reconstructions are moderately predictive of the presence of nonunion when both facets are not fused. As with any diagnostic test, surgeons need to interpret findings on fine-cut CT scans in conjunction with the patient's clinical presentation to determine when surgical intervention is necessary.

Key Points

- The kappa for interobserver agreement was moderate for evaluating facet fusions (0.42) and substantial for posterolateral fusions (0.62).
- A CT scan reading of either one or both posterolateral gutters fused or both facets fused are moderately predictive of a solid fusion on surgical exploration.
- Fine-cut CT scans with reconstructions are moderately predictive of the presence of nonunion when both facets are not fused.

References

1. Brodsky AE, Evan SK, Momtaz AK. Correlation of radiologic assessment of lumbar spine fusions with surgical exploration. *Spine* 1991;16(suppl):261–5.
2. Laasonen EM, Soini J. Low back pain after lumbar fusion: surgical and computed tomographic analysis. *Spine* 1989;14:210–3.
3. Landis RJ, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74.
4. Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature: III. How to use an article about a diagnostic test. B: what are the results and will they help me in caring for my patients? *JAMA* 1994;271:703–7.
5. Blumenthal SL, Gill K. Can lumbar spine radiographs accurately determine fusion in postoperative patients? Correlation of routine radiographs with second surgical look at lumbar fusions. *Spine* 1993;18:1186–9.
6. Larsen JM, Rimoldi RL, Capen DA, Nelson RW, et al. Assessment of pseudoarthrosis in pedicle screw fusion: A prospective study comparing plain radiographs, flexion/extension radiographs, CT scanning and Bone Scintigraphy with operative findings. *J Spinal Disord* 1996;9:117–20.
7. Kant AP, Daum WJ, Dean SM, et al. Evaluation of lumbar spine fusion: plain radiographs versus direct surgical exploration and observation. *Spine* 1995; 20:2313–7.
8. Morris J, Chafetz N, Baumrind S, et al. Stereophotogrammetry of the lumbar spine: a technique for the detection of pseudoarthrosis. *Spine* 1985;10: 368–75.
9. Slizofski WJ, Collier DB, Flatley TJ, et al. Painful pseudoarthrosis following lumbar spine fusion: detection by combined SPECT and planar bone scintigraphy. *Skeletal Radiol* 1987;16:136–41.
10. Lang P, Genant HK, Chafetz N, et al. Three-dimensional computed tomography and multiplanar reformations in the assessment of pseudoarthrosis in posterior lumbar fusion patients. *Spine* 1988;13:69–75.
11. Zinreich SJ, Long DM, Davis R, et al. Three-dimensional CT imaging in postsurgical 'failed back' syndrome. *J Comput Assist Tomogr* 1990;14: 574–80.
12. Chafetz N, Cann CE, Morris JM, et al. Pseudoarthrosis following lumbar fusion: detection by direct coronal CT scanning. *Radiology* 1987;162: 803–5.

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

ELSEVIER

The Spine Journal 7 (2007) 39–43

THE SPINE JOURNAL

Reliability and agreement between fine-cut CT scans and plain radiography in the evaluation of posterolateral fusions

Leah Y. Carreon, MD, MSc[a,*], Steven D. Glassman, MD[a,b], Mladen Djurasovic, MD[a,b]

[a]*Leatherman Spine Center, 210 East Gray Street, Suite 900, Louisville, KY 40202, USA*

[b]*Department of Orthopaedic Surgery, University of Louisville School of Medicine, 210 East Gray Street, Suite 1003, Louisville, KY 40202, USA*

Received 16 January 2006; accepted 5 April 2006

Abstract

BACKGROUND CONTEXT: Current imaging techniques used to evaluate fusion status after a posterolateral fusion such as radiographs, computed axial tomography (CT) scans, and tomograms are known to be inaccurate, with error rates estimated from 20% to 40%. Previous studies evaluated CT scans using 2–4-mm thick slices with limited reconstructions.

PURPOSE: The purpose of this study is to determine the intraobserver and interobserver agreement of plain radiographs and fine-cut (1-mm) CT scans with sagittal and coronal reconstructions in evaluating fusion status after instrumented posterolateral fusions. The correlation between radiographic evaluations and CT scan evaluations was also analyzed.

STUDY DESIGN/SETTING: Cross-sectional, blinded.

PATIENT SAMPLE: One-year radiographs and CT scans of 86 patients who had single-level instrumented posterolateral fusions.

OUTCOME MEASURES: Fusion grades based on previously published criteria were determined.

METHODS: Three spine surgeons graded the fusions of 86 patients who had single-level instrumented posterolateral fusions using 1-year postoperative flexion/extension lateral and anteroposterior radiographs, and fine-cut CT scans with sagittal and coronal reconstructions. The technique used to obtain the radiographs and the CT scans was the same in all cases. Two separate readings, 2 weeks apart, were done on each patient by each surgeon. The kappa coefficients for interobserver and intraobserver variability were determined.

RESULTS: The intraobserver agreement using CT scans to assess fusion status was moderate for both classification systems (Molinari=0.48, Glassman 0.47). The intraobserver agreement using X-rays to assess fusion status was fair for the Molinari classification (kappa=0.37) and moderate for the Glassman classification (kappa=0.43). The interobserver agreement using CT scans to assess fusion status was moderate for both classification systems (Molinari=0.48, Glassman 0.48). The interobserver agreement using X-rays to assess fusion status was fair for both classification systems (Molinari=0.24, Glassman 0.26). Observers agreed most often when the fusion was assessed as solid (Molinari k=0.61, Glassman k=0.63). The rating on the radiographs and CT scans agreed only 46% to 59% of the time.

CONCLUSIONS: Fine-cut CT scans with reconstructions have a considerably greater degree of interobserver and intraobserver agreement compared with flexion/extension and anteroposterior radiographs. Observers agree most often when the fusion is assessed as solid. Fusion evaluation based on radiographs agrees with CT scans only half the time. Future studies are needed to correlate the findings on fine-cut CT scans with surgical exploration. © 2007 Elsevier Inc. All rights reserved.

Keywords: CT scan; Radiographs; Spine fusion; Reliability; Interobserver variability; Kappa coefficient

FDA device/drug status: not applicable.

Research supported by a Research Grant from Norton Healthcare, Inc.

* Corresponding author. Leatherman Spine Center, 210 East Gray Street, Suite 900, Louisville, KY 40202. Tel.: (502) 584-7525; fax: (502) 584-6851.

E-mail address: lcarreon@spinemds.com (L.Y. Carreon)

1529-9430/06/$ – see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.spinee.2006.04.005

Introduction

Regardless of the surgical technique used to perform an arthrodesis, the problem of determining whether a fusion is solid or not remains. Clinically, determining the presence of a solid fusion becomes particularly important in patients who have undergone a lumbar fusion and have new or recurrent symptoms. Controversy regarding the overall efficacy of spinal fusion in treating lumbar degenerative conditions also highlights the need for a reliable and accurate noninvasive method to determine the status of a spinal fusion.

Several diagnostic imaging techniques have been used to assess the status of a spinal fusion. These include plain radiographs [1–3], bending films [2,4], stereophotogrammetry [5–8], X-ray polytomography [4,9], magnetic resonance imaging, and radionuclide imaging [2,6]. Various techniques of computed axial tomography (CT) [2,4,10], including two-dimensional multiplanar reformations and three-dimensional reconstructions [11,12], as well as direct coronal CT scanning [13] have also been used.

Previous studies that used only axial CT scans may have missed horizontal pseudoarthrosis [2,4]. Older CT scan techniques using 2–4-mm-thick axial slices to generate two-dimensional multiplanar reformations and three-dimensional reconstructions could have significantly limited the resolution and diagnostic information available. Current high-resolution CT scanners produce contiguous 1-mm-thick axial sections with a 1-mm table incrementation to optimize spatial resolution and to enhance the quality of computer-generated reformatted images. Contiguous axial images and contiguous reformatted sagittal and coronal images decrease the probability that bony bridging over a very limited area will be missed. Thinner sections also provide improved spatial resolution and improve the quality of reformatted images.

The purpose of this study is to determine the intraobserver and interobserver variability of plain radiographs and fine-cut (1-mm) CT scans with sagittal and coronal reconstructions in evaluating fusion status after a single-level instrumented posterolateral fusion. The correlation between radiographic evaluations and CT scan evaluations was also analyzed.

Methods

This was a cross-sectional, blinded study. One-year postoperative flexion/extension lateral and anteroposterior radiographs and fine-cut CT scans with sagittal and coronal reconstructions in 86 patients who had single-level instrumented posterolateral fusions were evaluated. All patients underwent a posterolateral fusion with titanium pedicle screws and rods (CD Horizon; Medtronic Sofamor Danek, Memphis, TN). The use of titanium minimizes artifacts created by metallic implants when evaluating CT scans. The fusion was done through a midline incision followed by complete exposure of the transverse processes, facets, and sacral ala when needed. These surfaces were decorticated, followed by an osteotomy of the facet joints and removal of the joint cartilage. Graft was then packed onto the decorticated surfaces.

Table 1
Molinari grading system for fusion

1	Definite	Solid trabeculated transverse process and facet fusion.
2	Probable	Thick fusion mass on one side. Difficult to visualize on other side
3	Probably not	Suspected lucency or defect in the fusion mass.
4	Not fused	Definite resorption of graft with fatigue of instrumentation.
5		Unable to assess.

The technique used to obtain the radiographs and the CT scans was the same in all cases. Anteroposterior radiographs were taken in the standing position. Patients who had a lumbosacral fusion had a Ferguson view taken. Flexion and extension radiographs were taken in the lateral recumbent position, with care taken to prevent rotation of the patient or obliquity of the X-ray beam or radiographic plate. Flexion and extension radiographs were evaluated, even when motion across the fusion mass was not considered, to lessen radiation exposure, as these views were the standard of care. CT scans were 1-mm-thick, continuous, nonoverlapping axial slices. The gantry was tilted to obtain scans parallel to the disc space and stayed constant throughout the scan. The field of view was made as small as possible but still encompassing the two vertebrae to include both transverse processes. Window and level settings were 2000/350 on the GE scanners (General Electric, Fairfield, CT) to optimize trabecular bone detail. All the films reviewed were hard copies and not computer-assisted filmless digital images.

Three spine surgeons graded the fusions using two previously published grading systems: Molinari et al. [14] (Table 1) and Glassman et al. [15] (Table 2). To facilitate statistical analysis, the numeric grading for the Glassman classification was reversed with 1 being solid fusion and 5 as no fusion. Presence of broken instrumentation, lucencies around the screws, and motion across the fusion mass were not considered in the evaluation of the images. Rigid fixation, which was used in all the cases in this study, could have prevented the motion normally associated with nonunions [16]. Two separate readings, 2 weeks apart, were done on each patient by each surgeon. The kappa coefficient for interobserver and intraobserver variability was

Table 2
Modified Glassman posterolateral fusion grading system*

1	Solid bilateral fusion
2	Solid unilateral fusion
3	Partial bilateral fusion
4	Partial unilateral fusion
5	No fusion

* The original grading system had 1 as No fusion and 5 as Solid bilateral fusion. To simplify statistical analysis, the numerical assignment was reversed.

Table 3
Consensus rating for radiographs

	Molinari		Glassman	
Rating	Assessment 1	Assessment 2	Assessment 1	Assessment 2
1	36	37	37	39
2	28	25	16	21
3	10	6	13	11
4	1	1	4	2
5	3	1	2	2
No consensus	8	16	14	11

determined. The strength of agreement was interpreted based on the classification of Landis and Koch [17].

Five-by-five contingency tables were constructed to present the agreement between the radiographic and CT scan evaluations.

Results

Table 3 presents the consensus data for each of the rating scales using radiographs for each assessment, and Table 4 presents the consensus data for each of the rating scales using CT scans. The consensus rating is defined as the agreement of at least two raters. Note that there are several cases in which no consensus could be reached. Table 5 summarizes the level of intraobserver agreement for each of the rating scales and diagnostic techniques, overall and by observer. The values presented are kappa statistics. The intraobserver agreement using CT scans to assess fusion status was moderate for both classification systems (Molinari=0.48, Glassman 0.47). The intraobserver agreement using X-rays to assess fusion status was fair for the Molinari classification (kappa=0.37) and moderate for the Glassman classification (kappa=0.43).

Table 6 shows the interobserver agreement kappa statistics for each of the rating scales and diagnostic techniques. A kappa statistic was calculated for each of the possible ratings on each scale (provided that the rating was given at least once). The interobserver agreement using CT scans to assess fusion status was moderate for both classification systems (Molinari=0.48, Glassman 0.48). The interobserver agreement using X-rays to assess fusion status was fair for both classification systems (Molinari=0.24, Glassman 0.26).

Table 4
Consensus rating for CT scans

	Molinari		Glassman	
Rating	Assessment 1	Assessment 2	Assessment 1	Assessment 2
1	47	46	48	48
2	18	23	14	17
3	10	9	12	10
4	4	0	2	2
5	0	0	2	2
No consensus	7	8	8	7

Table 5
Level of intraobserver agreement for each of the rating scales and diagnostic techniques, overall and by observer

	CT scans		X-rays	
Observer	Molinari	Glassman	Molinary	Glassman
1	0.42	0.43	0.36	0.41
2	0.39	0.43	0.45	0.51
3	0.61	0.55	0.28	0.37
Overall	0.47	0.48	0.37	0.43

The values presented are kappa statistics.

Comparing the interobserver agreement using CT scans to the agreement using X-rays, observers were more likely to agree with respect to their ratings using the CT scan rather than the X-rays. Further, the observers seem to agree most when solid fusion was present. Pair-wise kappa coefficients did not uncover a "disagreeable rater", ie, a situation in which two raters are generally in agreement while the third consistently disagreed with the two. Of note also is that no rater used the Molinari grade "Unable to assess" for the CT scan images. In contrast, 21 of 86 sets of radiographs had at least one observer grade it as "Unable to assess".

Table 7 summarizes the agreement between the radiographic and CT scan evaluation based on Molinari's classification. Cases in which no consensus could be reached for either the radiographic rating or CT scan rating were not included. Note that more than half of the cells are poorly populated. Table 8 summarizes the agreement between the radiographic and CT scan evaluation based on Glassman's classification. Again, more than half of the cells are poorly populated.

Discussion

Several studies have evaluated different diagnostic imaging techniques used to assess the status of a spinal fusion. The simplest, cheapest, and still probably the most commonly used imaging technique is the plain anteroposterior and lateral radiograph with or without bending films. Brodsky et al. [4], reporting on 175 patients who had posterolateral fusion requiring revision, demonstrated 64%

Table 6
Level of interobserver agreement for each of the rating scales and diagnostic techniques

	CT scans		X-rays	
Grade	Molinari	Glassman	Molinari	Glassman
1	0.61	0.63	0.48	0.46
2	0.39	0.49	0.16	0.18
3	0.40	0.40	0.06	0.16
4	0.31	0.03	0.10	0.02
5		0.32	0.01	0.15
Overall	0.48	0.48	0.24	0.26

Note that a kappa statistic is calculated for each of the possible ratings on each scale (provided that the rating was given at least once).

Table 7
5×5 Contingency table showing agreement between X-ray rating and CT scan rating using the Molinari classification*

	CT scan rating									
	Assessment 1					Assessment 2				
X-ray rating	Definite	Probable	Probably not	Not fused	Unable to assess	Definite	Probable	Probably not	Not fused	Unable to assess
Definite	25	6	2	0	0	24	10	2	0	0
Probable	11	4	5	1	0	12	7	3	0	0
Probably not	4	4	2	0	0	6	3	0	0	0
Not fused	0	0	0	0	0	0	0	0	0	0
Unable to assess	3	0	0	1	0	0	0	1	0	0
Totals	43	14	9	2	0	42	20	6	0	0

* Cases in which consensus was lacking wer excluded.

correlation between radiographs and surgical exploration in determining the status of fusion. Blumenthal and Gill [1] had four observers evaluate preoperative anteroposterior and lateral radiographs of 49 patients and compared them with findings on surgical exploration of the fusion and removal of the instrumentation. The overall correlation between radiographic and surgical findings was 69%, with a false positive rate of 42% (range, 0–75%) and a false negative rate 29% (range, 20–51%). The kappa coefficient for intraobserver agreement in the study by Blumenthal and Gill ranged from fair to good agreement beyond chance (kappa=0.42 to 0.72). This is similar to the present study showing fair to moderate agreement. Kant et al. [3] reported on a similar study with one observer and 75 patients and demonstrated only a 68% correlation between the radiographic and surgical findings. These three studies show the lack of accuracy of plain radiographs in assessing the status of the fusion.

After radiographs, CT scan is the next most frequently ordered diagnostic imaging modality used to evaluate the status of an osseous fusion. Brodsky et al. [4] reported a 57% correlation between fusion assessment using 6-mm axial slice CT scans and surgical exploration. Laasonen and Soini [10], using 6-mm CT scan with selective sagittal reconstructions, found an 80% correlation between findings on the CT scan and surgical exploration. One might expect to obtain a better correlation when CT scans are performed, as this imaging technique should provide the greatest amount of information in assessing osseous anatomy and the status of a bony fusion. However, both these studies were retrospective and the CT scans were not done specifically to assess the status of the spinal fusion. In these three reports, the thickness of the axial sections could have limited the resolution and diagnostic information available in these imaging studies, even when reformatting was performed.

Our current study evaluated CT scans with contiguous 1-mm-thick axial sections with a 1-mm table incrementation, and reformatted sagittal and coronal images were obtained. These CT scans were taken specifically to assess the fusion healing. The intraobserver and interobserver agreement (0.47–0.48) showed moderate agreement beyond chance. Although the CT scans had a higher degree of agreement in evaluating the fusion status compared with radiographs, it is still not substantial. The higher degree of agreement is also reflected in the lower number of cases in which there was a lack of consensus when CT scans were rated compared with the number of cases lacking consensus when radiographs were used. The kappa coefficient only reaches a substantial degree of interobserver agreement when definite fusion has occurred.

Pair-wise kappa coefficients showed that the three observers were generally in agreement. This may be due to the similar training background of the raters. All the raters were fellowship trained spine surgeons who have been in practice for several years. Interestingly, none of the raters

Table 8
5×5 Contingency table showing agreement between X-ray rating and CT scan rating using the Glassman classification*

	CT scan rating									
	Assessment 1					Assessment 2				
X-ray Rating	Solid bilateral	Solid unilateral	Partial bilateral	Partial unilateral	No fusion	Solid bilateral	Solid unilateral	Partial bilateral	Partial unilateral	No fusion
Solid biliateral	26	7	2	0	0	29	5	2	0	0
Solid unilateral	10	2	2	3	0	12	8	1	0	0
Partial bilateral	7	1	4	0	1	3	3	3	0	1
Partial unilateral	1	1	2	0	0	0	0	1	1	0
No fusion	0	1	0	0	0	0	0	1	0	0
Totals	44	12	10	3	1	44	16	8	1	1

* Cases in which conserisus was lacking were excluded.

graded "Unable to assess" on the Molinari grading when evaluating CT scans. The greater amount of information gleaned from a CT scan with two-dimensional multiplane reformatting compared with radiographs, as reported previously by Zinreich et al. [12], allows the rater to more confidently grade the fusion.

The rating on the radiographs and CT scans agreed only 46% of the time in both assessments using the Molinari classification, and 46% of the time during the first assessment and 59% of the time in the second assessment using the Glassman classification. Thus, fusion evaluation based on radiographs cannot accurately predict the fusion evaluation on CT scans.

Fine-cut CT scans with reconstructions have a considerably greater degree of interobserver and intraobserver agreement compared with anteroposterior and flexion-extension radiographs. Observers agree most when definite fusion has been achieved. This is problematic, as improved radiographic techniques would be most clinically relevant in cases where nonunion is suspected or fusion is uncertain. Fusion evaluation based on radiographs agrees with CT scans only half the time. To determine the accuracy of current CT scan techniques in evaluating the status of a fusion, a prospective study to correlate findings using current CT scan techniques to surgical exploration needs to be done.

References

[1] Blumenthal SL, Gill K. Can lumbar spine radiographs accurately determine fusion in postoperative patients? Correlation of routine radiographs with second surgical look at lumbar fusions. Spine 1993;18:1186–9.

[2] Larsen JM, Rimoldi RL, Capen DA, et al. Assessment of pseudoarthrosis in pedicle screw fusion: a prospective study comparing plain radiographs, flexion/extension radiographs, CT scanning and bone scintigraphy with operative findings. J Spinal Dis 1996;9:117–20.

[3] Kant AP, Daum WJ, Dean SM, Uchida T. Evaluation of lumbar spine fusion: plain radiographs versus direct surgical exploration and observation. Spine 1995;20:2313–7.

[4] Brodsky AE, Evan SK, Momtaz AK. Correlation of radiologic assessment of lumbar spine fusions with surgical exploration. Spine 1991;16:S261–5.

[5] Morris J, Chafetz N, Baumrind S, Genant H, Korn EL. Stereophotogrammetry of the lumbar spine: a technique for the detection of pseudoarthrosis. Spine 1985;10:368–75.

[6] Johnsson R, Axelsson P, Gunnarsson G, Stromqvist B. Stability of lumbar fusion with transpedicular fixation determined by roentgen stereophotogrammetric analysis. Spine 1999;24:687–90.

[7] Johnsson R, Selvik G, Stromqvist B, Sunden G. Mobility of the lower lumbar spine after posterolateral fusion determined by roentgen stereophotogrammetric analysis. Spine 1990;15:347–50.

[8] Olsson TH, Selvik G, Willner S. Kinematic analysis of spinal fusions. Invest Radiol 1976;11:202–9.

[9] Slizofski WJ, Collier DB, Flatley TJ, et al. Painful pseudoarthrosis following lumbar spine fusion: detection by combined SPECT and planar bone scintigraphy. Skeletal Radiol 1987;16:136–41.

[10] Laasonen EM, Soini J. Low back pain after lumbar fusion: surgical and computed tomographic analysis. Spine 1989;14:210–3.

[11] Lang P, Genant HK, Chafetz N, Steiger P, Morris JM. Three-dimensional computed tomography and multiplanar reformations in the assessment of pseudoarthrosis in posterior lumbar fusion patients. Spine 1988;13:69–75.

[12] Zinreich SJ, Long DM, Davis R, Quinn CB, McAfee PC, Wang H. Three-dimensional CT imaging in postsurgical "Failed Back" Syndrome. J Comput Assist Tomog 1990;14:574–80.

[13] Chafetz N, Cann CE, Morris JM, Steinbach LS, Goldberg HI, Ax L. Pseudoarthrosis following lumbar fusion: detection by direct coronal CT scanning. Radiology 1987;162:803–5.

[14] Molinari RW, Bridwell KH, Klepps SJ, Baldus C. Minimum 5 year follow-up of anterior column structural allografts in the thoracic and lumbar spine. Spine 1999;24:967–72.

[15] Glassman SD, Dimar JR, Carreon LY, Campbell MJ, Johnson JR, Puno RM. Initial fusion rates with rh-BMP-2 and HA-TCP/Collagen carrier in posterolateral fusion. Spine 2005;30:1694–8.

[16] Frymoyer JW, Hanley EN, Howe J, Kuhlman D, Materi RE. A comparison of radiographic findings in fusion and non-fusion patients ten or more years following lumbar disc surgery. Spine 1979;4:435–9.

[17] Landis RJ, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74.

LWWOnline | LOGOUT | eALERTS | PROFILE | CUSTOMER SUPPORT

MARCI DIPROFIL

LIPPINCOTT WILLIAMS & WILKINS

Spine An international journal for the study of the spine

Home Search Current Issue Archive

Return to the search results

ARTICLE LINKS:
Abstract | References (28) | View full-size inline images

Spine: Volume 22(24) 15 December 1997 pp 2807-2812

1997 Volvo Award Winner in Clinical Studies: Degenerative Lumbar Spondylolisthesis With Spinal Stenosis: A Prospective, Randomized Study Comparing Decompressive Laminectomy and Arthrodesis With and Without Spinal Instrumentation [Clinical Studies]

Fischgrund, Jeffrey S. MD*; Mackay, Michael MD*; Herkowitz, Harry N. MD*; Brower, Richard MD; Montgomery, David M. MD*; Kurz, Lawrence T. MD*

From the *William Beaumont Hospital, Department of Orthopaedic Surgery, Royal Oak, Michigan.

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

Acknowledgment date: July 8, 1997.

Acceptance date: July 8, 1997.

Device status category: 4.

Address reprint requests to: Jeffrey S. Fischgrund, MD; 16800 West Twelve Mile Road; Suite 100; Southfield, MI 48076-2176.

Abstract TOP

Study Design. This prospective study analyzed the influence of transpedicular instrumented on the operative treatment of patients with degenerative spondylolisthesis and spinal stenosis.

Objectives. To determine whether the addition of transpedicular instrumented improves the clinical outcome and fusion rate of patients undergoing posterolateral fusion after decompression for spinal stenosis with concomitant degenerative spondylolisthesis.

Summary of Background Data. Decompression is often necessary in the treatment of symptomatic patients who have degenerative spondylolisthesis and spinal stenosis. Results of recent studies demonstrated that outcomes are significantly improved if posterolateral arthrodesis is performed at the listhesed level. A meta-analysis of the literature concluded that adjunctive spinal instrumentation for this procedure can enhance the fusion rate, although the effect on clinical outcome remains uncertain.

Article Outline

- Abstract
- Materials and Methods
- Results
- Discussion
- References
- Citing Articles

Figures/Tables

- Figure 1
- Table 1
- Table 2

Methods. Seventy-six patients who had symptomatic spinal stenosis associated with degenerative lumbar spondylolisthesis were prospectively studied. All patients underwent posterior decompression with concomitant posterolateral intertransverse process arthrodesis. The patients were randomized to a segmental transpedicular instrumented or noninstrumented group.

Results. Sixty-seven patients were available for a 2-year follow-up. Clinical outcome was excellent or good in 76% of the patients in whom instrumentation was placed and in 85% of those in whom no instrumentation was placed ($P = 0.45$). Successful arthrodesis occurred in 82% of the instrumented cases *versus* 45% of the noninstrumented cases ($P = 0.0015$). Overall, successful fusion did not

influence patient outcome ($P = 0.435$).

Conclusions. In patients undergoing single-level posterolateral fusion for degenerative spondylolisthesis with spinal stenosis, the use of pedicle screws may lead to a higher fusion rate, but clinical outcome shows no improvement in pain in the back and lower limbs

Junghanns,[14] in 1931, was the first to describe degenerative lumbar spondylolisthesis. He defined the entity of lumbar vertebral spondylolisthesis without a pars interarticularis defect as "pseudospondylolisthesis." Newman[19] noted that listhesis of the vertebral body with an intact neural arch was usually caused by degenerative arthritis of the lumbar facet joints. Since these early descriptions, degenerative lumbar spondylolisthesis has been extensively studied, but operative management remains controversial. Surgical options reported include decompression,[3,4] decompression and arthrodesis,[8,10] and decompression and arthrodesis with spinal instrumentation.[28]

In the past, the recommended surgical procedure for this condition was decompressive lumbar laminectomy alone. However, in 1991, Herkowitz and Kurz[10] published a randomized prospective study comparing the results of decompressive lumbar laminectomy alone with lumbar laminectomy with posterolateral arthrodesis. Fifty patients were assigned alternatively to the two treatment groups and were evaluated for a mean of 3 years after surgery. The results of this study indicated that those patients who had concomitant arthrodesis had an statistically significant improvement in clinical outcome. Although pseudarthrosis was noted in nine patients (36%) in the arthrodesis group, all patients in whom pseudarthrosis developed had a good or an excellent result.

Recently, several surgeons[1,22,28] have advocated the addition of spinal instrumentation in the operative management of patients with degenerative spondylolisthesis and spinal stenosis. The theoretical advantage of instrumentation is postulated to be an increased fusion rate, decreased rehabilitation time,[9] and, most importantly, an improved patient outcome. However, the indications for the use of spinal instrumentation remain controversial.

To determine these indications, a randomized prospective study was performed, involving patients with degenerative spondylolisthesis a single level associated with lumbar spinal stenosis. This study compared the results of decompression and arthrodesis alone with those of decompression and arthrodesis combined with instrumentation at the level of the arthrodesis.

Materials and Methods TOP

Sixty-eight consecutive patients, who agreed to participate in a clinical study approved by the Human Investigational Committee at William Beaumont Hospital, Royal Oak, Michigan, are included in this report. All patients had a clinical diagnosis of degenerative spondylolisthesis and spinal stenosis. Most patients (94%) complained of back pain, which usually was aggravated by activity and relieved by rest. All patients had significant buttock and leg pain before the surgical procedure. The most common leg complaints were due to neurogenic claudication. Typically, patients complained of pain, numbness, tingling, weakness, cramping, or burning, beginning in the low back and buttocks and radiating into one or both legs after walking. All patients had undergone a trial of nonoperative treatment for at least 3 months before surgery. Nonoperative treatment included physical therapy (passive modalities and aerobic exercise) and nonsteroidal antiinflammatory drugs, if tolerated. Patients were recommended for surgery if they had failed nonoperative treatment and continued to have significant pain and/or significant daily activity restrictions due to neurogenic claudication or radicular pain.

All patients were noted on plain radiographs to have a single-level degenerative lumbar spondylolisthesis, and imaging studies (computed tomographic myelogram and/or magnetic resonance imaging) demonstrated spinal stenosis at the level of the spondylolisthesis. No patients had undergone prior lumbar spinal surgery.

The patients were assigned randomly to one of two treatment groups: Decompressive laminectomy and single level autogenous bilateral lateral intertransverse process arthrodesis, or decompressive laminectomy and single level bilateral lateral autogenous intertransverse process arthrodesis with transpedicular instrumentation. Randomization occurred at the time the decision was made to proceed with surgical intervention, by the withdrawal of a card from an envelope which indicated either instrumentation or no instrumentation. Randomization was performed by a medical assistant, not by the treating physician.

There were 55 women and 13 men. Seven men and 28 women had instrumentation placed, and 6 men and 27 women had an arthrodesis performed without instrumentation. The ages of the patients who had instrumentation ranged from 53 to 86 years (mean, 69 years) and those of the patients who did not have instrumentation, from 52 to 80 years (mean, 66 years). In the entire study, 5 patients were 80 years of age or older (4 instrumentation, 1 noninstrumentation). Seven patients were smokers (four instrumentation, three noninstrumentation).

The operation was performed at L4-L5 in 69 patients, at L3-L4 in 6 patients, and at L5-S1 in 1 patient. Informed, written consent was obtained from each participant. Before the operation, plain radiographs of the lumbosacral spine (including anteroposterior, lateral, left and right oblique, standing lateral, and standing flexion-extension lateral) were obtained for all patients and were repeated at the most recent follow-up evaluation. Preoperative and follow-up radiographic films were analyzed to determine the amount of

spondylolisthesis, in millimeters, on the lateral radiographs; the amount of sagittal motion, in millimeters, on the flexion-extension lateral radiographs;[7,18,25] and the total amount of angular motion, in degrees, between the adjacent vertebral end plates at the operative location seen on flexion-extension radiographs.

Arthrodesis was deemed successful if final follow-up radiographs demonstrated a continuity in the fusion mass between the cephalad and caudad transverse processes. Pseudarthrosis was determined to be present if there was no continuity in the fusion mass or if lateral flexion-extension radiographs demonstrated greater than 2° of angular motion between the adjacent end plates or greater than 2 mm of sagittal motion at the location of the spondylolisthesis. (Figure 1, A and C).

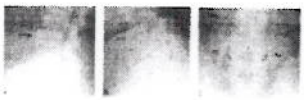


Figure 1. **A,** Two-year postoperative lateral extension radiograph demonstrating 8 mm of subluxation at L4-L5. **B,** Two-year postoperative lateral radiograph demonstrating an increase of the subluxation to 12 mm during flexion. **C,** Two-year postoperative anteroposterior radiograph demonstrating clefts (arrows) in the lateral fusion mass between L4 and L5.

Decompression of the central canal and nerve roots was performed by removing half of the cephalad and the caudad lamina of the involved vertebra, together with bilateral medial caudad and cephalad facetectomy. The technique of spinal arthrodesis was that described by MacNab and Dall[16] and by Wiltse[26] for a single-level bilateral intertransverse process arthrodesis. The outer table of the iliac crest was exposed through the same skin incision that was used for the decompression and arthrodesis. Strips of cortical-cancellous and cancellous bone were harvested from the outer and middle tables of the iliac crest and were placed across the transverse processes,[27,28] after decortication of the transverse processes with a bur or a rongeur.

Pedicle screws (VSP, Acromed, Cleveland, OH) were placed at the location of the spondylolisthesis, according to the method of West et al.[24] The point of insertion of the screw was at the junction of the middle of the transverse process and the superior facet. The cortical bone at the starting point was perforated with a bur, and a pedicle probe was used to locate the transpedicular canal and to create a channel through the pedicle into the vertebral body for subsequent placement of the screw. Under fluoroscopic visualization, a tap was then directed into the pedicle, followed by screw placement. The width of the pedicle in the axial plane had been determined before surgery. The largest screw that would fit within the pedicle was then inserted. Generally, the entry hole was tapped approximately 1 mm less than the screw length to afford better bone purchase by the screw. After the four screws were placed, two plates (Acromed) were bolted in position with lock nuts. Care was taken to assure that the plate did not extend proximally to the cephalad screw, to avoid impingement on the unfused superior facet.

Before the operation, all patients rated pain in the back and lower limbs (including that in the buttocks) on a visual analog scale, ranging from 0 (no pain) to 5 points (severe pain). Pain in the back was rated separately from that in the lower limbs. The scoring procedure was repeated at the final follow-up examination.

The operative results were rated as excellent, good, fair, or poor, as previously described.[10] The result was considered to be excellent if the patient resumed unrestricted activity and had near complete relief of pain in the back, lower limbs, or both. A good result indicated that there was occasional discomfort in the back or lower limbs, necessitating occasional nonnarcotic medication. Patients with a good result had significant improvement, compared with the preoperative condition, and had resumed unrestricted activity. A fair result was defined as intermittent discomfort in the back, lower limbs, or both; improvement compared with the preoperative condition; restriction of activities; and an occasional need for nonnarcotic medication. The patients who had a poor result had marked discomfort in the back, lower limbs, or both, necessitating nonnarcotic and occasional narcotic medication. The patients in this category noted no improvement compared with the preoperative condition and had significant restriction of activities.

All clinical and radiographic assessments were made by examiners other than the treating surgeons. Radiographs were independently examined by two orthopedic surgeons. If the reported fusion status differed between the examiners, the radiographs were reexamined and a consensus reached.

The clinical results of the operation and radiographic findings were analyzed with the use of Student's *t* test of independent samples, Mann-Whitney test, sign test, and Fisher's exact test.

The same postoperative treatment was used for both groups of patients. Walking was permitted on the first postoperative day and progressed during the first 4 to 6 weeks after surgery. Exercises on a stationary bike or in-water therapy began at 6 to 8 weeks, and exercises for flexion of the spine and strengthening of the abdominal muscles were added at 10 to 12 weeks. No brace or corset was used after surgery in either group. The duration of follow-up ranged from 2 to 3 years (mean follow-up, 28 months).

Results TOP

All patients who were randomized to the instrumented group had four pedicle screws successfully implanted at the time of surgery. There were no intraoperative findings that required withdrawal of a patient from the study, nor did any patient's random assignment

change at the time of surgery.

Clinical outcome, assessed according to relief of pain and increase in activity, was excellent or good in 78% of the patients who had instrumentation placed and in 85% of those patients who had no instrumentation placed. Statistical analysis revealed no significant difference in the results between the two groups (Fisher's exact test with midpoint *P* correction, $P = 0.45$).

In the instrumented group, 80% of the patients rated preoperative leg pain at 4 or 5 (average, 4), and 74% rated back pain at 4 or 5 (average, 4). At final follow-up, 64% rated leg pain at 0 or 1 (average, 1), and 58% rated back pain at 0 or 1 (average, 1). Statistical analysis showed significant reduction in leg pain ($P < 0.001$) and in back pain ($P < 0.001$).

There was also a significant reduction of pain in the noninstrumented group. Before surgery, 89% of the patients rated leg pain at 4 or 5 (average, 4); at final follow-up, 75% rated leg pain at 0 or 1 (average, 1). Initial back pain in this group was rated at 4 or 5 by 65% of the patients (average, 4); at final follow-up 53% reported pain at 0 or 1 (average, 2).

Successful arthrodesis occurred in 83% of the instrumented spines *versus* 45% of the noninstrumented ones, a statistically significant result ($P = 0.0015$). However, successful fusion was not predictive of successful patient outcome ($P = 0.435$).

Before surgery, both groups averaged 3 mm of sagittal motion (range, 2-18 mm) and 9° of angular motion (range, 0-11°) on lateral flexion and extension radiographs. The spondylolisthesis measured 8 mm in the instrumented group and 7 mm in the noninstrumented group, as demonstrated by neutral standing preoperative lateral radiographs. After surgery, spondylolisthesis decreased in the instrumented group to 6 mm, whereas sagittal and angular motion decreased to 1 mm and 1°, respectively. The noninstrumented group had no change in spondylolisthesis at final follow-up, with sagittal and angular motion decreasing to 2 mm and 5°, respectively (Table 1).

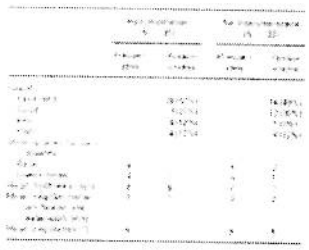


Table 1. Data on the 68 Patients

The significant, continued angular motion in the noninstrumented group reflected the relatively high nonfusion rate. To determine which, if any, variables contributed to pseudarthrosis, preoperative angular motion, spondylolisthesis, and sagittal motion were analyzed and related to fusion outcome. Combining both groups of patients (instrumented and noninstrumented), preoperative angulation averaged 8° in those patients who eventually had a successful fusion, compared with 11° in those in whom pseudarthrosis developed ($P = 0.066$; Student's *t* test of independent samples). Preoperative spondylolisthesis ($P = 0.28$) and sagittal motion ($P = 0.18$) were not statistically significant in fusion outcome (Table 2).

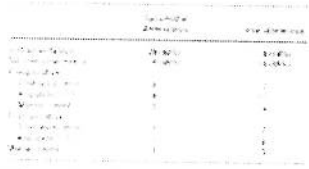


Table 2. Factors Affecting Fusion Rate

There were no new peripheral (lower motor neuron) neurologic deficits after surgery in either group. No patients required early hardware removal because of persistent radicular pain, and no postoperative infections developed. Of the eight patients with poor results, five underwent further lumbar surgery at least 1 year after the index procedure. Two patients (one instrumented, one noninstrumented) required decompressive lumbar laminectomy at a spinal location different from that of the original surgery. One patient had hardware removed for persistent low back pain, and solid fusion was confirmed during the second surgery. One patient in the noninstrumented group, with persistent low back pain and pseudarthrosis, had a second attempt at arthrodesis, this time with instrumentation. In one patient (instrumented group) recurrent stenosis and pseudarthrosis developed, requiring a second decompression, instrumentation, and arthrodesis. There was only one screw failure (S1) in an asymptomatic patient, with solid fusion seen on radiographic film and an excellent clinical outcome.

Seven original participants in the study could not be located or refused to return at the required 2-year follow-up (five instrumented, two noninstrumented), and one patient died (cause unrelated to the surgical procedure) before the 2-year review. In that these patients did not complete the study, their data are not included in this report.

Discussion TOP

The majority of patients with spinal stenosis and degenerative spondylolisthesis respond to nonoperative treatment. For the patients in whom this regimen fails to produce improvement, the goals of surgery are relief of pain and improvement in quality of life. Previously, surgical management of this condition consisted of decompressive lumbar laminectomy alone.4 However, recent studies3,9,10,20 have produced results substantiating the value of arthrodesis with decompressive laminectomy.

The critical issue regarding instrumentation after intertransverse process arthrodesis in this condition is not only whether the rate of fusion will increase, but whether clinical outcome will also be improved. It is true that the purpose of arthrodesis is to obtain solid fusion, but it is also true that a good clinical outcome can be achieved without solid bony fusion.10 Numerous studies2,23 have outlined the difficulty in determining fusion status from radiographs, and methods for evaluating the fusion mass vary widely. The only accurate method is visual inspection, which is usually not practical. In the current study, the fusion mass was evaluated as critically as possible, using plain radiographs-hence, the low reported fusion rate, which is contrary to the high clinical success rate. The increased cost and complication rate associated with spinal instrumentation should be weighed against the successful outcome demonstrated in this report.

A recent meta-analysis17 of the published literature on degenerative spondylolisthesis included 889 patients from 25 publications. Reported studies were classified into the following groups: posterior decompression without arthrodesis, posterior decompression with arthrodesis but without instrumentation, and posterior decompression with arthrodesis and pedicle instrumentation.

Evaluation of the clinical results in the group undergoing decompression without arthrodesis revealed that 69%1,3,4,6,8,10,12,13,15,20,21 of patients had a satisfactory outcome. Progressive slipping after decompression was noted in most reports. Addition of arthrodesis to the decompression increased the satisfactory outcome to 90%,1,8,10,15,20,28 and 86% achieved solid fusion (range, 30-100%).

In this meta-analysis, five studies were included that described decompression with intertransverse process arthrodesis and instrumentation.1,5,11,23,28 There was no statistically significant difference in fusion rate ($P = 0.08$) between the group without instrumentation and the group with pedicle screws. Although the fusion rate was higher with instrumentation (93% *versus* 86%) the clinical outcome was better in the noninstrumented group (90% *versus* 86%).

The current series is the largest prospectively randomized study reporting on the use of pedicle screws for one diagnosis. Fusion rate was markedly increased in the instrumented group; however, there was no statistically significant difference in clinical outcome between the two groups. These conclusions are in agreement with those reported by other surgeons.1,28 Although pseudarthrosis developed in 55% of the noninstrumented patients, the clinical result was excellent or good in 15 of these 18 patients (83%). Radiographic fusion status did not affect clinical outcome. These results are in agreement with those obtained by Herkowitz and Kurz10 and may be related to the development of a fibrous fusion that provides sufficient structural support to prevent progressive spondylolisthesis.

In an attempt to identify those patients who were more likely to have pseudarthrosis, preoperative radiographic findings were analyzed. The only variable that approached statistical significance was preoperative angular motion at the location of the spondylolisthesis. In the 44 patients in whom successful fusion was achieved, the preoperative angulation averaged 8°, whereas angulation in the 24 patients with nonfusion averaged 11° before surgery ($P = 0.066$).

Prior prospective studies evaluating the use of pedicle screws in patients with spinal stenosis and degenerative spondylolisthesis were compared with the results reported here. Zdeblick28 reported on 124 patients undergoing lumbar or lumbosacral fusions for five degenerative conditions. The patients were randomized to one of three groups: Group 1, posterolateral fusion, using autogenous bone graft only; Group 2, autogenous posterolateral fusion, supplemented by a semirigid pedicle screw-plate fixation system; and Group 3, posterolateral autogenous fusion with a rigid pedicle screw-rod fixation system. Overall, the rigid pedicle fixation group had a significantly higher percentage of successful fusions (95%) than did the noninstrumented group (65%). Additionally, better clinical results were seen in the instrumented group (95% excellent or good) compared with those in the noninstrumented group (71% excellent or good).

A second study by Bridwell et al,1 reported on 44 patients with degenerative spondylolisthesis who underwent surgery, primarily for spinal stenosis. Patients were classified into one of three groups: Group 1, no arthrodesis performed; Group 2, Posterolateral arthrodesis without instrumentation; and Group 3, posterolateral arthrodesis with instrumentation. If excessive motion (more than 10° of angular motion or 3 mm of translational motion) at the slip location was noted on preoperative radiograph, the patient was not randomized but was automatically assigned to receive instrumentation. Results were an 87% fusion rate in the instrumented cases *versus* a 30% rate in noninstrumented cases. Functional status was improved in 83% of those patients receiving instrumentation. In

contrast with previous results, there was no significant difference in clinical outcome between Groups 1 and 2 (33% *vs.* 30% successful clinical outcome, respectively). The 30% clinical success noted in the fused and noninstrumented group is markedly lower than the 90% satisfactory outcome reported in the meta-analysis review of the literature. No explanation for this is apparent from the study data presented.

In summary, the results of the current study demonstrate that transpedicular instrumentation improves the fusion rate, after posterolateral fusion for patients with degenerative spondylolisthesis. However, clinical outcome assessed in terms of relief of pain and increase in activity is unchanged whether or not instrumentation is used.

References TOP

1. Bridwell KH, Sedgewick TA, O'Brien MF, Lenke LG, Baldus C. The role of fusion and instrumentation in the treatment of degenerative spondylolisthesis with spinal stenosis. J Spinal Disord 1993;6:461-72.
[Context Link]

2. Brodsky A, Kovalsky E, Khalil M. Correlation of radiologic assessment of lumbar spine fusions with surgical exploration. Spine 1991;16:261-5.
[Context Link]

3. Caputy A, Luessenhop A. Long-term evaluation of decompressive surgery for degenerative lumbar stenosis. J Neurosurg 1992;77:669-76.
[Medline Link] [Context Link]

4. Cauchoix J, Benoist M, Chassaing V. Degenerative spondylolisthesis. Clin Orthop 1976;115:123-9.
[Context Link]

5. Chang K, McAfee P. Degenerative spondylolisthesis and degenerative scoliosis treated with a combination segmental rod-plate and transpedicular screw instrumentation system: A preliminary report. J Spinal Dis 1989;1:247-56.
[Context Link]

6. Dall B, Rowe D. Degenerative spondylolisthesis: Its surgical management. Spine 1985;10:669-72.
[Context Link]

7. Dupuis PR, Yong-Hing Ken, Cassidy JD, Kirkaldy-Willis WH. Radiologic diagnosis of degenerative lumbar spinal instability. Spine 1985;10:262-76.
[Medline Link] [Context Link]

8. Feffer H, Weisel S, Cuckler J, Rothman R. Degenerative spondylolisthesis: To fuse or not to fuse. Spine 1985;10:286-9.
[Context Link]

9. Herkowitz HN. Degenerative lumbar spondylolisthesis. Spine 1996;20:1084-90.
[Context Link]

10. Herkowitz HN, Kurz LT. Degenerative lumbar spondylolisthesis with spinal stenosis. J Bone Joint Surg [Am] 1991;73:802-7.
[Medline Link] [Context Link]

11. Hirabayashi S, Kiyoshi K, Swaki T. Cotrel-Dubousset pedicle screw system for various spinal disorders. Spine 1991;16:1298-304.
[Medline Link] [Context Link]

12. Herron L, Trippi A. L4-L5 Degenerative spondylolisthesis: The results of treatment by decompressive laminectomy without fusion. Spine 1989;14:534-8.
[Medline Link] [Context Link]

13. Johnsson K, Redlund-Johnell I, et al. Preoperative and Postoperative instability in lumbar spinal stenosis. Spine 1989;14:591-3.
[Medline Link] [Context Link]

14. Junghanns H. Spondylolisthesen ohne Spalt im Zwischergelenkstuclz ("Pseudospondylolisthen"). Arch Orthop Unfallchirurgie 1931;29:118-27.
[Context Link]

15. Lombardi J, Wiltse L, et al. Treatment of degenerative spondylolisthesis. Spine 1985;10:821-7.
[Medline Link] [Context Link]

16. MacNab I, Dall D. The blood supply of the lumbar spine and its application to the technique of intertransverse lumbar fusion. J Bone Joint Surg [Br] 1971;53:628-38.
[Medline Link] [Context Link]

17. Mardjetko SM, Connolly PJ, Shott S. Degenerative lumbar spondylolisthesis: A meta-analysis of the literature 1970-1993. Spine 1994;10:2256S-2265S.
[Context Link]

18. Nasia RJ. Surgical management of lumbar spinal stenosis. Spine 1987;12:809-18.
[Context Link]

19. Newman P. Stenosis of the lumbar spine in spondylolisthesis. Clin Orthop 1976;115:116-21.

[Fulltext Link] [Medline Link] [CrossRef] [Context Link]

20. Postacchini F, Cinotti G. Bone regrowth after surgical decompression for lumbar spinal stenosis. J Bone Joint Surg [Br] 1992;74:862-9.
[Medline Link] [Context Link]

21. Sanderson P, Wood P. Surgery for lumbar spinal stenosis in old people. J Bone Joint Surg [Br] 1993;75:393-7.
[Medline Link] [Context Link]

22. Simmons E, Capicotto W. Posterior transpedicular zielke instrumentation of lumbar spine. Clin Orthop 1988;236:180-90.
[Fulltext Link] [Medline Link] [CrossRef] [Context Link]

23. Steffee A, Brantigan J. The variable screw placement spinal fixation system. Spine 1993;18:1160-72.
[Context Link]

24. West JL, Bradford DS, Ogilvie JW. Results of spinal arthrodesis with pedicle screw-plate fixation. J Bone Joint Surg [Am] 1991;73:1179-84.
[Medline Link] [Context Link]

25. White AA, III, Panjabi MM, Pusner I, Edwards WT, Hayes WC. Spinal stability: Evaluation and treatment. In: Instructional Course Lectures, The American Academy of Orthopaedic Surgeons. Vol. 30. St. Louis: CV Mosby, 1981:457-83.
[Context Link]

26. Wiltse LL, Bateman JG, Hutchinson RH, Nelson WE. The Paraspinal sacrospinalis-splitting approach to the lumbar spine. J Bone Joint Surg [Am] 1988;50:919-26.
[Context Link]

27. Wisneski RJ, Rothman RH. Posterior intertransverse process fusion: Indications, pathomechanics and results. In: White AH, Rothman RH, Ray CD, eds. Lumbar Spine Surgery: Techniques and Complications. St. Louis: CV Mosby, 1987:2721-2745.
[Context Link]

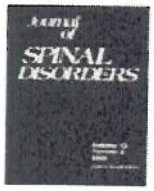
28. Zdeblick TA. A prospective, randomized study of lumbar fusion. Spine 1993;18:983-91.
[Context Link]

Keywords:
degenerative spondylolisthesis; lumbar stenosis; posterolateral fusion; transpedicular instrumentation

© Lippincott-Raven Publishers

Citing Articles TOP

Use of Recombinant Human Bone Morphogenetic Protein-2 as an Adjunct in Posterolateral Lumbar Spine Fusion: A Prospective CT-Scan Analysis at One and Two Years.
Journal of **Spinal** Disorders & Techniques. 19(6):416-423, August 2006.
Singh, Kern MD *; Smucker, Joseph D. MD +; Boden, Scott D. MD ++
[Abstract] [Fulltext] [PDF (581 K)]



Biomechanical Comparison of Anatomic Trajectory Pedicle Screw versus Injectable Calcium Sulfate Graft-Augmented Pedicle Screw for Salvage in Cadaveric Thoracic Bone.
Journal of **Spinal** Disorders & Techniques. 19(4):286-291, June 2006.
Derincek, Alihan MD; Wu, Chunhui PhD; Mehbod, Amir MD; Transfeldt, Ensor E. MD
[Abstract] [Fulltext] [PDF (228 K)]

Dynamic Degenerative Lumbar Spondylolisthesis: Diagnosis With Axial Loaded Magnetic Resonance Imaging.
Spine. 31(10):E298-E301, May 1, 2006.
Jayakumar, Prakash MBBS, BSc (Hons) *; Nnadi, Colin FRCS (Eng) *; Saifuddin, Asif MRCP, FRCR +++[S]; MacSweeney, Emer FRCR [S]; Casey, Adrian FRCS *
[Abstract] [Fulltext] [PDF (379 K)]

Graft Resorption With the Use of Bone Morphogenetic Protein: Lessons From Anterior Lumbar Interbody Fusion Using Femoral Ring Allografts and

Recombinant Human Bone Morphogenetic Protein-2.
Spine. 31(10):E277-E284, May 1, 2006.
Pradhan, Ben B. MD, MSE *; Bae, Hyun W. MD *; Dawson, Edgar G. MD *; Patel, Vikas V. MA, MD +; Delamarter, Rick B. MD *
[Abstract] [Fulltext] [PDF (957 K)]

Point of View: Dynamic Stabilization in Addition to Decompression for Lumbar Spinal Stenosis With Degenerative Spondylolisthesis.
Spine. 31(4):450, February 15, 2006.
Sengupta, Dilip K. MD
[Fulltext] [PDF (55 K)]

Dynamic Stabilization in Addition to Decompression for Lumbar Spinal Stenosis with Degenerative Spondylolisthesis.
Spine. 31(4):442-449, February 15, 2006.
Schnake, Klaus John MD *; Schaeren, Stefan MD +; Jeanneret, Bernard MD +
[Abstract] [Fulltext] [PDF (569 K)]

Comparison of OP-1 Putty (rhBMP-7) to Iliac Crest Autograft for Posterolateral Lumbar Arthrodesis: A Minimum 2-Year Follow-up Pilot Study.
Spine. 30(24):2709-2716, December 15, 2005.
Vaccaro, Alexander R. MD *; Anderson, D Greg MD +; Patel, Tushar MD ++; Fischgrund, Jeffrey MD [S]; Truumees, Eeric MD [//]; Herkowitz, Harry N. MD [P]; Phillips, Frank MD #; Hilibrand, Alan MD **; Albert, Todd J. MD **; Wetzel, Todd MD ++; McCulloch, John A. MD ++++
[Abstract] [Fulltext] [PDF (421 K)]

Surgical Treatment for the Painful Motion Segment: Matching Technology With the Indications: Posterior Lumbar Fusion.
Spine. Painful Motion Segment. 30(16S) Supplement:S44-S51, August 15, 2005.
Polly, David W. Jr MD; Santos, Edward R. G. MD; Mehbod, Amir A. MD
[Abstract] [Fulltext] [PDF (1.11 M)]

Current Treatment Strategies for the Painful Lumbar Motion Segment: Posterolateral Fusion Versus Interbody Fusion.
Spine. Painful Motion Segment. 30(16S) Supplement:S33-S43, August 15, 2005.
Wang, Jeremy C. MD *; Mummaneni, Praveen V. MD dagger;; Haid, Regis W. MD *
[Abstract] [Fulltext] [PDF (1.18 M)]

Lower Dose of rhBMP-2 Achieves Spine Fusion When Combined With an Osteoconductive Bulking Agent in Non-human Primates.
Spine. 30(10):1127-1133, May 15, 2005.
Barnes, Bryan MD *; Boden, Scott D. MD +; Louis-Ugbo, John MD +; Tomak, Patrick R. MD *; Park, Jin-Soo MD +; Park, Moon-Soo MD +; Minamide, Akihito MD +
[Abstract] [Fulltext] [PDF (1.02 M)]

Experimental Anterior Lumbar Interbody Fusion With an Osteoinductive Bovine Bone Collagen Extract.
Spine. 30(8):890-896, April 15, 2005.
Li, Haisheng MD, PhD; Zou, Xuenong MD, PhD; Woo, Charlotte; Ding, Ming MD, PhD; Lind, Martin MD, DMSc; Bunger, Cody MD, DMSc
[Abstract] [Fulltext] [PDF (1.23 M)]

Degenerative Spondylolisthesis: Review of Current Trends and Controversies.

Spine. 30(6S) Supplement:S71-S81, March 15, 2005.
Sengupta, Dilip K. MD *; Herkowitz, Harry N. MD +
[Abstract] [Fulltext] [PDF (838 K)]

Adult Low-Grade Acquired Spondylolytic Spondylolisthesis: Evaluation and Management.
Spine. 30(6S) Supplement:S35-S41, March 15, 2005.
Kwon, Brian K. MD, PhD, FRCSC *; Albert, Todd J. MD +
[Abstract] [Fulltext] [PDF (494 K)]

Clinical Experience With the Dynesys Semirigid Fixation System for the Lumbar Spine: Surgical and Patient-Oriented Outcome in 50 Cases After an Average of 2 Years.
Spine. 30(3):324-331, February 1, 2005.
Grob, Dieter MD *; Benini, Arnoldo MD *; Junge, Astrid PhD *; Mannion, Anne F. PhD *+
[Abstract] [Fulltext] [PDF (430 K)]

The Influence of Subdiagnosis on Radiographic and Clinical Outcomes After Lumbar Fusion for Degenerative Disc Disorders: An Analysis of the Literature From Two Decades.
Spine. 30(2):227-234, January 15, 2005.
Bono, Christopher M. MD *; Lee, Casey K. MD +
[Abstract] [Fulltext] [PDF (317 K)]

A Pilot Study Evaluating the Safety and Efficacy of OP-1 Putty (rhBMP-7) as a Replacement for Iliac Crest Autograft in Posterolateral Lumbar Arthrodesis for Degenerative Spondylolisthesis.
Spine. 29(17):1885-1892, September 1, 2004.
Vaccaro, Alexander R. MD *; Patel, Tushar MD +; Fischgrund, Jeffrey MD ++; Anderson, D Greg MD [S]; Truumees, Eeric [P]; Herkowitz, Harry N. MD ++; Phillips, Frank MD ++; Hilibrand, Alan MD *; Albert, Todd J. MD *; Wetzel, Todd MD ++; McCulloch, John A. MD [S][S]
[Abstract] [Fulltext] [PDF (615 K)]

Lumbar Synovial Cysts: A Review of Diagnosis, Surgical Management, and Outcome Assessment.
Journal of Spinal Disorders & Techniques. 17(4):321-325, August 2004.
Epstein, Nancy E MD
[Abstract] [Fulltext] [PDF (184 K)]

Prospective Assessment of Outcomes Improvement Following Fusion for Low Back Pain.
Journal of Spinal Disorders & Techniques. 17(3):183-188, June 2004.
Robertson, Peter A. MD, FRACS; Jackson, Suzanne A. FRCS, FRACS
[Abstract] [Fulltext] [PDF (78 K)]

Point of View.
Spine. 29(7):733-734, April 1, 2004.
Katz, Jeffrey N. MD, MS
[Fulltext] [PDF (497 K)]

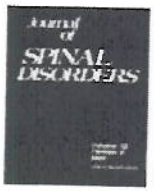
A 10-Year Follow-up Evaluation of Lumbar Spine Fusion With Pedicle Screw Fixation.
Spine. 28(13):1390-1395, July 1, 2003.
Glaser, John MD *; Stanley, Mark MD +; Sayre, Hutha RN *; Woody, Joyce *; Found, Ernest MD *; Spratt, Kevin PhD *
[Abstract] [Fulltext] [PDF (362 K)]

Bone graft alternatives in spinal fusion surgery.
Current Opinion in Orthopedics. 14(3):127-137, June 2003.
Kim, David H. MD *; Jenis, Louis MD *; Berta, Scott C. MD +; Vaccaro, Alexander R. MD +
[Abstract] [Fulltext] [PDF (750 K)]

Circumferential Lumbar Spinal Fusion With Brantigan Cage Versus Posterolateral Fusion With Titanium Cotrel-Dubousset Instrumentation: A Prospective, Randomized Clinical Study of 146 Patients.
Spine. 27(23):2674-2683, December 1, 2002.
Christensen, Finn B. MD, PhD; Hansen, Ebbe S. MD DMSc; Eiskjaer, Soren P. MD; Hoy, Kristian MD; Helmig, Peter MD, PhD; Neumann, Pavel MD, PhD; Niedermann, Bent MD; Bunger, Cody E. MD, DMSc
[Abstract] [Fulltext] [PDF (1.12 M)]

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. 27(23):2662-2673, December 1, 2002.
Boden, Scott D. MD; Kang, James MD; Sandhu, Harvinder MD; Heller, John G. MD
[Abstract] [Fulltext] [PDF (2.86 M)]

Chronic Low Back Pain and Fusion: A Comparison of Three Surgical Techniques: A Prospective Multicenter Randomized Study From the Swedish Lumbar Spine Study Group.
Spine. 27(11):1131-1141, June 1, 2002.
Fritzell, Peter MD *; Hagg, Olle MD +; Wessberg, Per MD +; Nordwall, Anders MD, PhD ++; the Swedish Lumbar Spine Study Group
[Abstract] [Fulltext] [PDF (456 K)]



The Influence of Lumbar Lordosis on Spinal Fusion and Functional Outcome After Posterolateral Spinal Fusion With and Without Pedicle Screw Instrumentation.
Journal of Spinal Disorders & Techniques. 15(3):187-192, June 2002.
Korsgaard, Marianne; Christensen, Finn Bjarke; Thomsen, Karsten; Hansen, Ebbe Stender; Bunger, Cody
[Abstract] [Fulltext] [PDF (333 K)]

Lumbosacral Stability of Consolidated Anteroposterior Fusion After Instrumentation Removal Determined by Roentgen Stereophotogrammetric Analysis and Direct Surgical Exploration.
Spine. 27(3):269-274, February 1, 2002.
Pape, Dietrich MD ,*; Fritsch, Ekkehard MD ,*; Kelm, Jens MD ,*; Muller, Katja *; Georg, Thomas PhD +; Kohn, Dieter *; Adam, Frank MD *
[Abstract] [Fulltext] [PDF (322 K)]

Prosthetic Disc Replacement: The Future?
Clinical Orthopaedics & Related Research. 394:139-145, January 2002.
Bao, Qi-Bin PhD *; Yuan, Hansen A. MD **
[Abstract] [Fulltext] [PDF (98 K)]

A Double-Blind Study of Capacitively Coupled Electrical Stimulation as an Adjunct to Lumbar Spinal Fusions.
Spine. 24(13):1349, July 1, 1999.
Goodwin, Charles B. MD *; Brighton, Carl T. MD +; Guyer, Richard D. MD ++; Johnson, John R. MD [S]; Light, Kenneth I. MD [//]; Yuan, Hansen A. MD [P]
[Abstract] [Fulltext]

Copyright © 2006, Lippincott Williams & Wilkins. All rights reserved.
Published by Lippincott Williams & Wilkins.
Copyright/Disclaimer Notice • Privacy Policy
Subscribe to RSS feed
utrdc-pt01
Release 4.4.2

J Neurosurg Spine 7:486–495, 2007

Use of osteogenic protein-1 in patients at high risk for spinal pseudarthrosis: a prospective cohort study assessing safety, health-related quality of life, and radiographic fusion

Invited submission from the Joint Section on Disorders of the Spine and Peripheral Nerves, March 2007

Julio C. Furlan, M.D., M.B.A., M.Sc., Ph.D.,[1,2] Richard G. Perrin, M.D.,[1,2] Preneshlin V. Govender, M.D.,[1,2] Yuriy Petrenko, M.D.,[1,2] Eric M. Massicotte, M.D., M.Sc., F.R.C.S.C.,[1,2] Yoga R. Rampersaud, M.D., F.R.C.S.C.,[3] Stephen Lewis, M.D., F.R.C.S.C.,[3] and Michael G. Fehlings, M.D., Ph.D., F.R.C.S.C.[1,2]

[1]Division of Neurosurgery, Department of Surgery, University of Toronto; [2]Spinal Program, Krembil Neuroscience Centre, Toronto Western Hospital, University Health Network; and [3]Division of Orthopedic Surgery, Department of Surgery, University of Toronto, Toronto Western Hospital, University Health Network, Toronto, Canada

Object. The capability of osteogenic protein (OP)–1 to induce bone formation has led to an increasing interest in its use in fusion surgery. This prospective study examines the safety and efficacy of OP-1 use in patients considered to be at a high risk for developing pseudarthrosis following reconstructive spinal surgery.

Methods. Outcome measures included documentation of adverse events, radiographic evaluation of fusion by an independent musculoskeletal radiologist blinded to treatment, the Oswestry Disability Index (ODI), and the 36-Item Short Form Health Survey (SF-36). The health-related quality of life (HRQOL) assessments (ODI and SF-36) were given at baseline and at 3, 6, 12, 18, and 24 months after the surgical OP-1 implant.

Results. The study consisted of 17 male and 13 female patients, with a mean age of 53 years (range 20–77 years). Fourteen patients underwent operations for cervical disease, and 16 for lumbar disease, with a median postoperative follow-up of 24 months (range 13–46 months). There were significant improvements in the physical health (from 28.7 ± 1.5 to 34.2 ± 3; $p = 0.025$) and mental health (from 43.7 ± 2 to 47.5 ± 3.1; $p = 0.015$) summary scores on the SF-36. The mean postoperative ODI score at 6, 9, 12, and 18 months was significantly lower than the baseline ODI score, after taking into consideration a 10-point measurement error ($p = 0.0003$, $p = 0.003$, $p = 0.004$, and $p = 0.032$, respectively). At 24 months, however, the differences in ODI scores were no longer significant. Of the 30 patients, 24 (80%) were deemed to have a solid fusion. There were no allergic reactions to OP-1 and no symptomatic postoperative hematomas.

Conclusions. Our results suggest that the use of OP-1 is safe and may contribute to high fusion rates, as demonstrated by radiographs, reduced levels of disability, and improved HRQOL in patients considered to be at a high risk for developing a nonunion after spinal reconstructive surgery. *(DOI: 10.3171/SPI-07/09/486)*

Key Words • **bone morphogenetic protein–7** • **osteogenic protein-1** • **Oswestry Disability Index** • **radiographic fusion** • **spinal pseudarthrosis** • **36-Item Short Form Health Survey**

Bone morphogenetic and osteogenetic proteins are multifunctional growth factors that belong to the transforming growth factor–β superfamily.[3] Although osteogenetic proteins are primarily considered osteogenic factors, further investigations have shown that these proteins are also essential for embryogenesis and organogenesis, and that they have pleiotropic roles in cell growth, differentiation, migration, and apoptosis.[3,11] The rhOP-1, also known as bone morphogenetic protein–7, has been documented as a potential treatment alternative for different diseases, including bone disease, stroke, inflammatory bowel disease, prostate cancer, and chronic renal disease.[5,20,22,30,40]

Abbreviations used in this paper: BMI = body mass index; CT = computed tomography; HRQOL = health-related quality of life; MR = magnetic resonance; ODI = Oswestry Disability Index; OP = osteogenic protein; rhOP-1 = recombinant human OP-1; SF-36 = 36-Item Short Form Health Survey.

More explicitly, OP-1 plays an important role in bone formation by inducing differentiation of pluripotent mesenchymal cells into active osteoblasts.[32,34] Results from animal studies[8,10,14,21] have shown that use of OP-1 can induce a stable, mature, posterolateral spinal fusion mass more rapidly than an autologous bone graft, and the resulting fusion mass may be biomechanically stiffer in the early stages (up to 3 months) of healing.[10,14]

Osteogenic protein receptor signaling plays a key role in normal postnatal bone formation, and several experimental studies in animals have demonstrated the capability of OPs to enhance or replace autograft bone for spinal arthrodesis.[2,4] The OPs promote osteoblast proliferation and induce differentiation of progenitor cells into osteoblasts that form bone.[28] Upregulation of OP is observed at fracture sites during the early stage of fracture repair.[28] The bone formation–inducing capacity of OPs has also been demonstrated in both bone defect models and in vivo heterotopic sites.[23] Although OP-1 has been reported to overcome the adverse effects of nicotine on fusion in animals, this growth factor was unable to overcome the inhibitory effects of estrogen deficiency on spinal fusion.[15,23,25] In addition, the rate of fracture healing was greater in 3-month-old rats than in 18-month-old rats, but OP-1 can effectively induce fracture repair in both young and old animals.[17] Therefore, gender and age are the two major potential confounders that need to be controlled for in clinical studies.

Despite the fact that the effects of OPs have been extensively studied, there are still unanswered questions regarding the safety and beneficial effects of OPs in patients who undergo spinal surgery. At our institution, rhOP-1 has been used in selected adult patients at high risk for pseudarthrosis following a posterior spinal fusion. The objectives of this pilot study were fourfold: 1) longitudinally examine HRQOL and disability in a cohort of patients at high risk for spinal pseudarthrosis who undergo spine surgery involving an implant of rhOP-1 putty; 2) evaluate the beneficial effects of an OP-1 implant in terms of radiological fusion; 3) assess the potential influence of age, sex, BMI, ethnicity, number of fused levels, and level of spine disease on long-term outcome in these patients after the rhOP-1 implant; and 4) evaluate the safety of an rhOP-1 implant in this group of patients.

Clinical Material and Methods

Study Population

The Canadian Health Protection Branch and the Ontario Ministry of Health and Long-Term Care approved the use of rhOP-1 in our patients based on a compassionate use protocol. The Research Ethics Board of the University Health Network approved the research protocol for this study. All patients who agreed to participate signed a consent form. Inclusion criteria consisted of adult patients at a high risk for pseudarthrosis following a posterior spinal fusion who had a baseline assessment and a minimum postoperative follow-up of 12 months. The population at risk for a spinal nonunion was defined as patients with connective tissue disorders, individuals with a history of major medical comorbidities that could adversely affect bone healing, patients receiving medications that negatively affect bone healing, patients with a history of previous nonunion fusions, and/or patients with limited availability or poor quality of autogenous bone graft.

Intervention With rhOP-1 Putty Implant

The rhOP-1 implant consisted of 3.5 mg of lyophilized rhOP-1 and a carrier consisting of 1 g of Type 1 bovine bone collagen (Stryker Biotech). The dry powder was reconstituted in 2.5 ml of saline to form a putty just before surgical implantation. The amount of rhOP-1 used in this study protocol was determined according to information from previous animal and human studies.[7–9,13]

After induction of general anesthesia and administration of prophylactic antibiotics, all patients underwent a routine posterior midline approach to the spine. Standard decompressions were performed as necessary to decompress the neural elements, and the required posterior instrumentation was performed to obtain and/or maintain spinal alignment and stability. Autologous bone was harvested from the iliac crest and/or posterior elements of the decompressed levels. The bone was decorticated using a high-speed drill, and the autologous bone was thoroughly mixed with the rhOP-1 putty and placed bilaterally in the posterolateral gutters. In the majority of the patients, one vial of rhOP-1 was used on either side of the spine, so that a total of two vials were given to those patients.

Safety Assessment

Adverse events in the study population were tracked using a prospective database, previously described by our team elsewhere.[27] In addition, heterotopic ossification was tracked postoperatively using plain radiographs, CT scans, and MR imaging. All patients underwent postoperative CT scans and MR imaging to document the adequacy of decompression and development of fusion, and to exclude peridural heterotopic ossification.

Outcome Measures

The outcome measures included an HRQOL assessment using the SF-36 (US version 1.0) and evaluation of the degree of disability using the ODI.[12,38] These outcome measurements were performed at baseline (the preoperative evaluation) and at 3, 6, 12, 18, and 24 months after the rhOP-1 putty was surgically implanted. The minimum clinically important differences that ultimately reflect measurement error were established based on previously reported data for the SF-36 (seven points in each domain) and for the ODI (10 points).[16,24,26]

In addition to the self-assessment for HRQOL and degree of disability for all patients, this longitudinal study included an evaluation of spinal stability based on the static and dynamic plain radiographs in those patients who underwent cervical spine fusion between 3 and 6 months after rhOP-1 putty was surgically implanted. The status of instrumentation was also assessed using radiographs. Radiographic instability was defined as translation of larger than 2 mm and/or angulation of more than 5° on postoperative flexion/extension radiographs. A solid fusion was defined as evidence of bridging bone on radiographic evaluation. Therefore, a successful radiological outcome was established if no radiographic instability was observed

along with radiological evidence of intact instrumentation and evidence of an osseous union. An independent musculoskeletal radiologist (Dr. David Salonen) reviewed all radiographs in a blind manner.

Statistical Analysis

Comparisons between the baseline and follow-up time with regard to SF-36 assessments and ODI scores were performed using the Student paired t-test. The potential risk factors for instrumentation-related problems (such as age, sex, BMI, ethnicity, level of spine disease, and number of fused levels) after the rhOP-1 was implanted were analyzed using the two-sided Mann–Whitney U-test and the two-sided Fisher exact test. All data analysis was performed using SAS statistical software (version 8.02, SAS Institute, Inc.). A probability value less than 0.05 was considered statistically significant.

Results

There were 30 patients included in the study, with a mean age of 53.1 years and a median age of 52 years (range 20–77 years). The majority of patients were Caucasian with high educational levels (Table 1). Lumbar or lumbosacral fusions were performed slightly more frequently than cervical or occipitocervical fusions (Table 1). Nearly half of this cohort had had at least one previous pseudarthrosis following spinal surgery. The mean postoperative follow-up duration was 23.8 months (median 24 months, range 13–46 months).

TABLE 1

Summary of the characteristics of the 30 patients in the study

Characteristic	Number of Cases (%)
sex	
male	17 (56.7)
female	13 (43.3)
ethnicity	
Caucasian	27 (90)
African-American	1 (3.3)
East Indian	1 (3.3)
Southern Asian	1 (3.3)
educational level	
less than high school	1 (3.3)
graduated from high school	5 (16.7)
some college education	9 (30)
graduated from college	8 (26.7)
postgraduate school or degree	7 (23.3)
surgical level	
lumbar or lumbosacral fusion	16 (53)
cervical or occipitocervical fusion	14 (47)
risk factors for pseudarthrosis*	
previous nonfusion	14
rheumatoid arthritis	6
systemic lupus erithematosus	2
Maroteaux–Lamy syndrome	1
ankylosing spondylitis	1
use of steroids, immunosuppressors	8
heavy smoking	6
osteopenia or osteoporosis	4

* The total number of patients exceeds 30 because many patients had more than one risk factor.

The SF-36 Assessment

Although there were no significant differences between baseline and postoperative assessment scores with regard to the general health domain on the SF-36 (42 ± 2.5 compared with 42.1 ± 2.5, respectively; p = 0.94), the physical function domain score improved from 30 ± 2 to 35.3 ± 2.5 (p = 0.021), and the bodily pain domain score increased from 31.9 ± 1.4 to 39.7 ± 2 (p = 0.0002). There was also a trend toward an improvement in the physical role domain score, from 31.5 ± 1.6 to 35.5 ± 1.9 (p = 0.063). This trend contributed to an improvement in the physical health summary score, from 28.7 ± 1.5 to 34.2 ± 3 (p = 0.025; Fig. 1A), even after taking into consideration a 7-point measurement error. An improvement in the physical health summary score of at least 7 points was observed in 10 (33.3%) of 30 patients.

Additionally, the vitality domain score significantly increased from 41 ± 1.7 to 44.7 ± 2.3 (p = 0.042), the social function domain score significantly improved from 33.4 ± 2.2 to 41 ± 2.3 (p = 0.0001), and the emotional role domain score significantly increased from 36.7 ± 2.6 to 42.3 ± 2.6 (p = 0.054). There was a trend toward a higher mental health domain score at the last postoperative evaluation (46.8 ± 2.2) in comparison with the baseline mental health domain score (42.6 ± 1.9; p = 0.084). This trend contributed to a significant improvement in the mental health summary score, from 43.7 ± 2 to 47.5 ± 3.1 (p = 0.015; Fig. 1A). Eleven (36.7%) of the 30 patients showed a minimum improvement of 7 points in the mental health summary score.

The physical health summary score significantly improved at 6 months (p = 0.0005), 12 months (p = 0.0007), and 18 months (p = 0.012) after surgery (Fig. 1B). There was also a trend for an improvement in the physical health summary score at 9 months after surgery (p = 0.085). The baseline physical health summary score, however, did not significantly differ from the postoperative assessments at 3 months (p = 0.358) or at 24 months (p = 0.205).

Although the mental health summary score did not significantly change at 3 months after surgery (p = 0.149), there was a significant improvement in the mental health summary score at 6 months (p = 0.011), 9 months (p = 0.0002), 12 months (p = 0.0009), 18 months (p = 0.01), and 24 months (p = 0.034) after surgery (Fig. 1B).

The ODI Scores

Although the overall difference between the post- and preoperative ODI scores did not reach significance after taking into consideration a 10-point measurement error (52 ± 3.2% compared with 39 ± 4.6%, respectively; p = 0.41; Fig. 2A), a reduction in ODI scores of at least 10 points was observed in 18 (60%) of 30 patients. As illustrated in Fig. 2A, only the social life section ($p < 0.009$) of the ODI showed a significant improvement after applying the correction for measurement error. There was also a trend toward a postoperative decrease in pain intensity score compared with the baseline score (p = 0.094). The last postoperative ODI score, however, did not significantly differ from the baseline ODI score with regard to per-

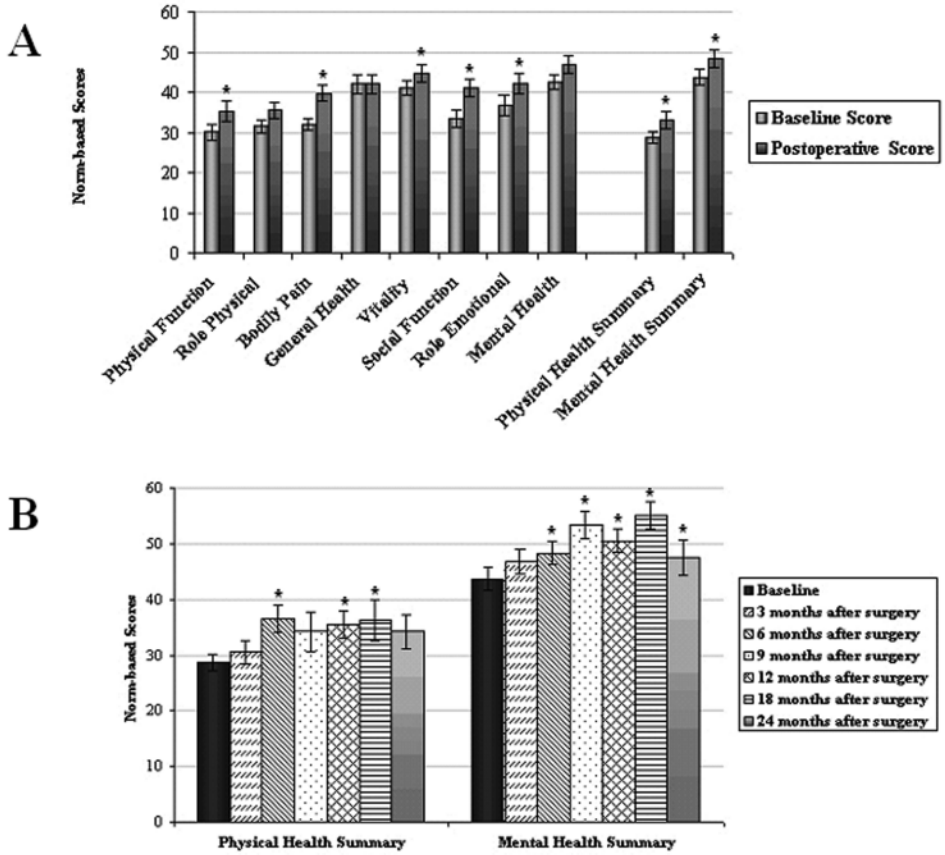


FIG. 1. Bar graphs of norm-based SF-36 scores comparing preoperative (baseline) and postoperative assessments of physical and mental health summary scores and their domains (A), and physical and mental summary scores over time (B). *Asterisks* indicate significant differences ($p < 0.05$) even after the inclusion of a seven-point measurement error in comparisons with baseline values.

sonal care ($p = 0.307$), lifting ($p = 0.804$), walking ($p = 0.603$), sitting ($p = 0.511$), standing ($p = 0.595$), sleeping ($p = 0.89$), sex life ($p = 0.659$), or traveling ($p = 0.797$) after taking into consideration a 10-point measurement error.

The mean postoperative ODI score at 6, 9, 12, and 18 months was significantly lower than the baseline ODI score ($p = 0.0003$, $p = 0.003$, $p = 0.004$, and $p = 0.032$, respectively) after taking into consideration a 10-point measurement error (Fig. 2B). The mean postoperative ODI score at 3 and 24 months, however, did not significantly differ from the baseline ODI score ($p = 0.323$ and $p = 0.204$, respectively) after applying the correction for measurement error (Fig. 2B).

Radiographic Assessment

After a median postoperative follow-up of 24 months, 24 (80%) of 30 patients were considered to have a solid fusion, which was defined as bridging bone, intact hardware, and an absence of motion on flexion/extension dynamic radiographs, as shown in illustrative cases of cervical fusion (Fig. 3) and lumbar fusion (Fig. 4). Although no patient demonstrated radiographic instability in terms of excessive angulation or translation on flexion/extension radiographs, six patients developed postoperative instrumentation-related problems, four of whom required a surgical revision (Table 2). Two patients (Cases 2 and 3) did not require reoperation (Table 2).

Potential Risk Factors for Instrumentation Problems

Although patients with solid fusions were apparently younger than patients with instrumentation-related problems after the surgical implantation of OP-1, there were no significant differences between these patient groups with regard to age (Table 3). In terms of sex, ethnicity, BMI, and level of disease, patients with solid fusion after rhOP-1 implantation did not significantly differ from patients who underwent rhOP-1 implantation but later developed instrumentation problems (Table 3). The number of fused levels, however, was significantly lower among patients with solid fusion after rhOP-1 implantation in comparison with patients who developed instrumentation problems (Table 3).

Safety of the OP-1 Implant

There was no evidence of systemic toxicity as defined by signs of anaphylaxis. One patient who underwent a successful occipitocervical fusion exhibited an asymptomatic linear opacification in the soft tissues, which likely represents heterotopic ossification (Fig. 5). All patients underwent MR imaging between 6 months and 1 year of follow-up. No patient showed evidence of peridural ossification.

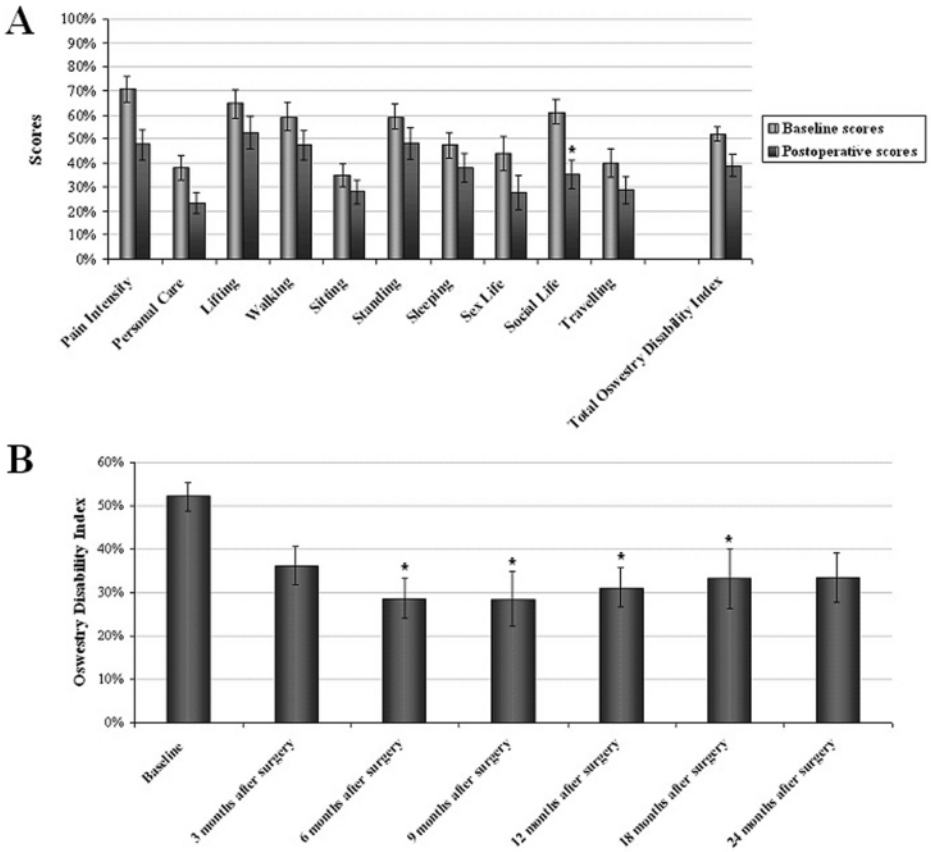


FIG. 2. Bar graphs of the ODI component scores comparing preoperative (baseline) with postoperative assessment (A), and ODI assessment scores over time (B). *Asterisks* indicate significant differences ($p < 0.05$) after the inclusion of a 10-point measurement error in comparisons with baseline values.

No patient in this series sustained neurological worsening. There were two superficial wound infections that responded to debridement and antibiotic administration. In addition, two other patients had thromboembolic disease (one with thrombosis of the aorta, and the other with a pulmonary thromboembolism) after their reoperation for revision of instrumentation failure, and both responded to anticoagulation therapy.

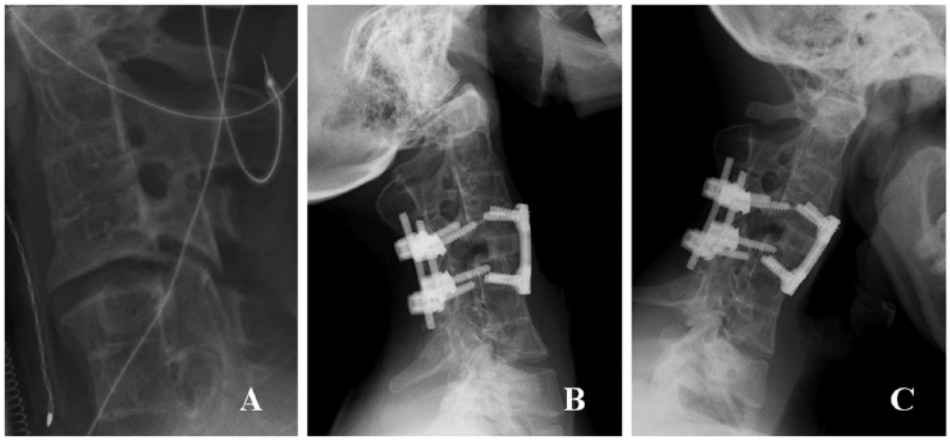


FIG. 3. Preoperative (A) and postoperative (B and C) radiographs obtained in a 24-year-old woman with a history of Klippel–Feil syndrome, juvenile rheumatoid arthritis, and immunosuppressant use. A: Image showing the development of C3–4 instability and transient quadriplegia after a diving injury. This injury was managed using combined anterior/posterior instrumentation as a staged procedure, resulting in solid fusion. The rhOP-1 implant was used to supplement the local autograft posteriorly. B and C: The absence of motion on extension (B) and flexion (C) dynamic radiographs was demonstrated at 36 months after spine fusion. This patient has a mild, asymptomatic C1–2 instability that has not required surgical intervention to date.

FIG. 4. Preoperative axial (A) and mid sagittal (B) CT scans and postoperative radiographs (C and D) obtained in a 69-year-old man with L4–5 spondylolisthesis and intractable neurogenic bladder due to spinal stenosis who underwent lumbar surgical fusion using bone grafting and adjunctive rhOP-1. Evidence for a solid fusion without instrumentation-related problems was demonstrated on the 27-month follow-up radiographs (C and D).

Discussion

The results of our study suggest that the adjunctive use of an rhOP-1 putty implant can improve HRQOL and reduce the degree of disability in patients who undergo spine surgery and are at high risk for spinal pseudarthrosis. In addition, a high rate of radiographic fusion was observed in this traditionally challenging cohort of patients. This potential benefit of rhOP-1 implantation appears to be independent of age, sex, ethnicity, BMI, and level of disease. Importantly, there were no apparent adverse effects related to the use of the OP-1 implant in this patient population.

Assessing HRQOL

The SF-36 is a generic measure of a patient's perception of his or her overall health state, which includes a physical component score (more heavily weighted for pain, physical function, and physical role function) and a mental component score (more heavily weighted for mental health, general health, and vitality).[1,39] The SF-36 is commonly used in studies in the orthopedic literature, including in several randomized controlled trials of spine surgery.[1] Although Riddle and Stratford[29] suggested that the SF-36 tends to be less sensitive to changes in specific orthopedic disorders such as spine diseases, a large multicenter study[37] has shown that this patient-based questionnaire may provide a sufficient measure of health status and patient function.

In a randomized clinical trial comparing the use of an rhOP-1 putty implant with an iliac crest autograft for posterolateral lumbar arthrodesis, Vaccaro and associates[36] found no significant differences in SF-36 scores at 12 months between the rhOP-1 treatment group (24 patients) and the autograft treatment group (12 patients). In a subsequent article, Vaccaro and colleagues[33] also reported similar results in terms of SF-36 score improvement in both

TABLE 2
Description of the instrumentation-related problems in this patient series

Case No.	Age (yrs), Sex	Preop Diagnosis	Description of Instrumentation Failure at Long-Term Follow-Up	Reoperation
1	75, M	C-1 arch fracture & posteriorly displaced C-2 dens fracture; nonunited odontoid fracture w/ some retrolisthesis of facet joint at C1–2; avascular necrosis of odontoid fracture (previous history of nonunions after various osseous injuries)	recurrence of myelopathy w/ imaging evidence of erosion of distal lamina at C3–4 secondary to occipito-cervical fixation hardware	revision of distal hardware
2	65, F	advanced spondylolisthesis, osteporosis, & severe lumbar kyphosis; previous attempt at fusion	broken screw unilaterally w/ some asymmetry of overall alignment but no progression	not required
3	67, M	central cord syndrome w/ atlantoaxial instability, spinal cord injury, & compressive myelopathy	fracture of transarticular screws bilaterally, but no evidence of instability in terms of excessive angulation or translation on flexion/extension radiographs	not required
4	62, F	pseudarthrosis w/ broken rods at L3–4 & L4–5 (kyphoscoliosis)	broken rod unilaterally	revision of hardware
5	57, F	scoliosis w/ loose hardware & pseudarthrosis (history of previous operation)	evidence of some loosening of proximal screw in T-12 (possible pseudarthrosis)	revision of hardware
6	46, M	spinal cord injury secondary to midthoracic fracture (20 yrs ago), syringomyelia, paraplegia at T-6 level, anxiety, significant narcotic requirement, & Charcot spine	failure of distal fixation	revision of hardware

TABLE 3

*Comparison between patients with solid fusions and patients with instrumentation-related problems after surgical rhOP-1 implant**

Characteristic	Solid-Fusion Group	Instrumentation-Related Problem Group	p Value
age (yrs)			
range	21–66	46–77	
mean ± SEM	50.7 ± 2.7	62.7 ± 4.8	0.102
sex (%)			
male	58.3	50	1
female	41.7	50	
ethnicity (%)			
Caucasian	87.5	100	1
non-Caucasian	12.5	0	
level of disease (%)			
lumbar/lumbosacral	50	66.7	0.657
cervical/occipitocervical	50	33.3	
no. of fused levels			
range	2–18	4–18	
mean (median)	6 (4)	10 (9)	0.027
BMI (kg/m^2)			
mean ± SEM	26.5 ± 1.2	23.4 ± 1.3	0.15

* SEM = standard error of the mean.

treatment groups at 24 months after surgery. In both studies, an increase in the SF-36 scores over time was documented even though no statistical comparison was provided between the baseline assessment and each follow-up evaluation.[33,36] The lack of information in those studies precludes more precise comparisons with our results. In addition, the patient population of those studies was different from our group of patients with regard to preexisting medical comorbidities, even though the baseline norm-based mental and physical health scores in the previous studies (28.8 and 47.0, respectively) are apparently similar to those in our group of patients (28.7 and 43.7, respectively).

In our study, most of the SF-36 domains in the baseline assessment significantly improved after spine surgery combined with the use of the rhOP-1 putty implant. Moreover, both postoperative physical and mental health summary scores showed significant improvement beyond the 7-point measurement error in comparison with baseline assessment scores. Whereas the physical health summary score significantly improved after surgical treatment at all time points after 6 months (but not at 24 months), the postoperative mental health summary score consistently showed significant improvement from 6 to 24 months. The lack of significant differences between baseline and postoperative physical health summary scores at 24 months should be carefully examined due to the relatively small number of patients (15) who answered the SF-36 questionnaires at this particular time point after surgery. Another possible explanation for this negative result would be the potential progression and impact of the other medical and musculoskeletal comorbidities within this patient population.[31]

Assessing the Degree of Disability

The ODI was included as a more specific outcome measure to assess low-back pain and patient functioning.[12] The ODI is a relatively reliable and valid measurement tool that has been considered to be one of the core set of measures for back pain.[1,12]

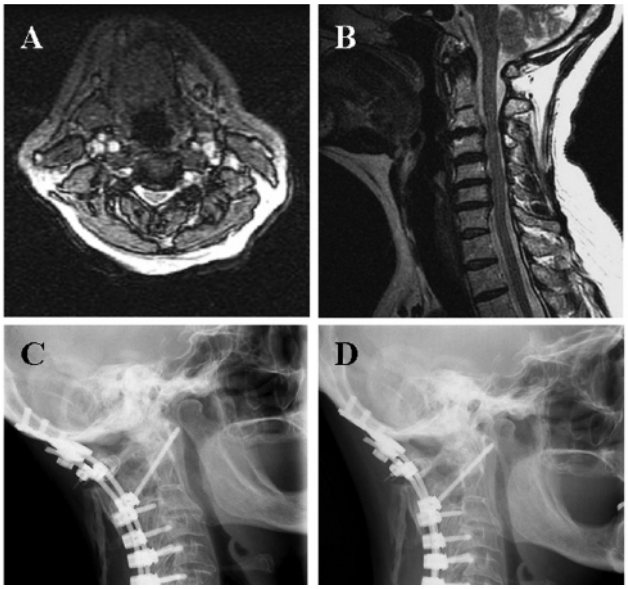


FIG. 5. Preoperative MR images (A and B) and postoperative radiographs (C and D) obtained in a 60-year-old woman with heterotopic ossification of soft tissues after surgical cervical fusion and reconstruction using bone grafting and rhOP-1. This patient developed atlantoaxial subluxation, Klippel–Feil deformity, C2–3 subaxial stable kyphosis, and C3–4 subaxial spondylosis with stenosis, as shown in the axial (A) and T2-weighted sagittal (B) images. The radiographic assessment performed 21 months after spine surgery (C and D) revealed a solid fusion and no instrumentation-related problems.

The ODI was commonly used as an outcome measure in previous studies of the safety and efficacy of the rhOP-1 putty implant. In a pilot study, Vaccaro and coworkers[35] reported that nine (75%) of 12 patients who received the rhOP-1 putty implant as an adjuvant to an iliac crest autograft in posterolateral lumbar fusions obtained at least a 20% improvement in their preoperative ODI score. In a randomized clinical trial, Vaccaro et al.[36] showed that 18 (86%) of 21 patients who received an rhOP-1 putty implant and 8 (73%) of 11 patients who received an autograft had at least a 20-point reduction in their baseline ODI score 12 months after posterolateral lumbar arthrodesis for degenerative spondylolisthesis. Subsequently, Vaccaro and colleagues[33] reported a 20-point reduction in the preoperative ODI score in 17 (85%) of 20 patients receiving an rhOP-1 implant and in 7 (64%) of 11 patients receiving an autograft 24 months after surgery to correct degenerative spondylolisthesis. At both time points, there were no significant differences between the rhOP-1 treatment group and the autograft treatment group with regard to clinical success, as defined by a 20% reduction in the baseline ODI score.[33,36] Although the ODI score in the rhOP-1 treatment group appeared to be lower than the score in the autograft treatment group at all time points, no statistical analysis was reported in terms of significant differences between the two groups at each time point.[33]

In another prospective randomized trial, Kanayama and associates[19] compared the use of an OP-1 putty implant with local autograft and ceramic bone substitute (hydroxyapatite/tricalcium phosphatate biphasic ceramic granules) in 20 patients with degenerative spondylolisthesis who un-

derwent posterolateral lumbar fusion. Although the ODI score was significantly decreased in both patient groups from 6 to 12 months after surgery, there were no statistically significant differences between the OP-1 and autograft groups.

In our study, there was no significant difference between the baseline and the last postoperative ODI assessment, even though a clinically significant difference between both assessments was observed in 60% of our patients. After comparing the baseline and postoperative ODI scores over time, the results suggest that at least a 10-point reduction in the ODI was attained at 6, 9, 12, and 18 months after surgical treatment. Because of some loss of follow-up between 18 and 24 months, the statistical power to detect a difference between baseline ODI and the final follow-up at 24 months was lower at this time point. Hence, although the ODI scores remained constant after 1 year, the lack of significance at 24 months most likely reflects a Type II error. Comparisons with previous studies might be inappropriate due to differences in preexisting medical comorbidities that could be factors in the baseline ODI scores of our group of patients (52 ± 3.2%) compared with the baseline ODI scores in two other studies (41 and 36.1%).[19,35]

Assessing Radiographic Fusion

Whereas investigators in several animal studies suggest that use of OP-1 may be superior to the use of an autologous bone graft in quickly achieving a solid fusion mass, this effect has not been demonstrated in clinical studies. Johnsson and coworkers[18] reported that there were no significant differences between radiostereometric and radiographic fusion results in 20 patients undergoing posterolateral fusion between L-5 and S-1 without instrumentation, randomized to receive either OP-1 alone or an autologous graft. Vaccaro and associates[35] demonstrated radiographically successful fusion in six (55%) of 11 patients undergoing single-level intertransverse fusion without instrumentation, using rhOP-1 putty and an autologous bone graft for degenerative spondylolisthesis; however, this rate did not significantly differ from a historical fusion rate of 45% for autologous bone graft alone. Vaccaro and coworkers[33] conducted a randomized clinical trial of posterolateral lumbar arthrodesis, in which, 24 months after surgery, the radiographic fusion rate in the groups of patients who received the rhOP-1 putty implant (55%) did not significantly differ from the rate obtained in the control group (40%) of patients receiving an iliac crest autograft. Using similar radiographic criteria, Kanayama and associates[19] observed that seven (78%) of nine patients who received the OP-1 putty implant and nine (90%) of 10 patients who received a local autograft and ceramic bone substitute had radiographic fusion after a minimum 12-month follow-up following posterolateral lumbar fusion for L3–4 or L4–5 degenerative spondylolisthesis.

Because our patient series included a heterogeneous group of patients, one could anticipate an increased risk of nonunion after posterolateral fusion with instrumentation. Based on our clinical experience, all patients had at least one risk factor for nonunion, and many had more than one risk factor. Almost half of the patients had previously experienced at least one pseudarthrosis, and four patients had more than one previous nonunion at the surgical site. Given the particularities of this cohort, it is encouraging that all 30 patients had radiographically stable fusions, even though a surgical revision for instrumentation-related problems was required in four patients (13.3%).

Potential Risk Factors for Instrumentation Problems

Investigators in preclinical studies have noted the potential effects of age and sex on bone spinal fusion.[17,23] Studying age-related changes in cartilage endogenous OP-1 of normal adult individuals, Chubinskaya et al.[6] observed a significant reduction (more than fourfold) in the OP-1 mRNA expression and protein levels with aging of normal adult cartilage. Their data suggest that OP-1 could serve as a repair factor for joint disease or aging.

Our study, for the first time, examined whether age, sex, ethnicity, BMI, and level of disease affect solid fusion as assessed using radiography in patients who received an rhOP-1 implant. In univariate analyses, none of those potential risk factors was found to significantly affect the radiographic fusion rate in this cohort of patients, but further investigation using a larger cohort of patients is needed to validate our preliminary results. Of note is the observation that a greater number of fused levels was associated with a risk for developing instrumentation problems in our cohort.

Safety of the OP-1 Implant

Previous investigational clinical studies have demonstrated no local or systemic adverse events related to the rhOP-1 putty implant.[33,35,36] In addition, Vaccaro and colleagues[36] showed that there were no significant differences regarding complication rates between the rhOP-1 treatment group and the autograft treatment group in their randomized clinical trial of posterolateral lumbar fusion for degenerative spondylolisthesis. The results of our study also indicate that use of the rhOP-1 putty implant carries no additional risk for adverse events in the treatment of patients at high risk for spinal pseudarthrosis.

Conclusions

This study contributes to the growing body of evidence that rhOP-1 is safe for surgical use in the clinical arena. The use of an rhOP-1 implant may be a contributing factor to the elevated rate of radiographic fusion, improved HRQOL, and reduced degree of disability after spinal fusion that was observed in this group of patients at a high risk for developing pseudarthrosis.

One may also speculate that the rhOP-1 implant is able to counteract factors that adversely affect bone fusion, and, therefore, its use may be superior to an autologous bone graft alone in patients at high risk for developing pseudarthrosis. The lack of either a control group or comparable historical data for this unique patient group, however, precluded us from providing a complete assessment of the independent affect of rhOP-1 on fusion. We hypothesize that rhOP-1 use does not confer additional benefit when bone fusion is likely to occur in a normal fashion; however, in patients with one or more risk factors for impaired bone healing, the adjuvant use of rhOP-1 may negate these effects, allowing adequate bone fusion to occur. A prospec-

tive, controlled trial to test this hypothesis will be conducted in the future at our institution.

Disclaimer

None of the authors has a financial interest in OP-1 or Stryker Biotech.

Acknowledgment

We gratefully acknowledge the contribution of Dr. David Salonen.

References

1. Beaton DE, Schemitsch E: Measures of health-related quality of life and physical function. **Clin Orthop Relat Res 413:**90–105, 2003
2. Carlisle E, Fischgrund JS: Bone morphogenetic proteins for spinal fusion. **Spine J 5 (6 Suppl):**240S–249S, 2005
3. Celeste AJ, Iannazzi JA, Taylor RC, Hewick RM, Rosen V, Wang EA, et al: Identification of transforming growth factor beta family members present in bone-inductive protein purified from bovine bone. **Proc Natl Acad Sci U S A 87:**9843–9847, 1990
4. Chen D, Zhao M, Mundy GR: Bone morphogenetic proteins. **Growth Factors 22:**233–241, 2004
5. Chou J, Harvey BK, Chang CF, Shen H, Morales M, Wang Y: Neuroregenerative effects of BMP7 after stroke in rats. **J Neurol Sci 240:**21–29, 2006
6. Chubinskaya S, Kumar B, Merrihew C, Heretis K, Rueger DC, Kuettner KE: Age-related changes in cartilage endogenous osteogenic protein-1 (OP-1). **Biochim Biophys Acta 1588:** 126–134, 2002
7. Cook SD, Baffes GC, Wolfe MW, Sampath TK, Rueger DC, Whitecloud TS III: The effect of recombinant human osteogenic protein-1 on healing of large segmental bone defects. **J Bone Joint Surg Am 76:**827–838, 1994
8. Cook SD, Dalton JE, Tan EH, Whitecloud TS III, Rueger DC: In vivo evaluation of recombinant human osteogenic protein (rhOP-1) implants as a bone graft substitute for spinal fusions. **Spine 19:**1655–1663, 1994
9. Cook SD, Wolfe MW, Salkeld SL, Rueger DC: Effect of recombinant human osteogenic protein-1 on healing of segmental defects in non-human primates. **J Bone Joint Surg Am 77:** 734–750, 1995
10. Cunningham BW, Shimamoto N, Sefter JC, Dmitriev AE, Orbegoso CM, McCarthy EF, et al: Osseointegration of autograft versus osteogenic protein-1 in posterolateral spinal arthrodesis: emphasis on the comparative mechanisms of bone induction. **Spine J 2:**11–24, 2002
11. Ducy P, Karsenty G: The family of bone morphogenetic proteins. **Kidney Int 57:**2207–2214, 2000
12. Fairbank JC, Couper J, Davies JB, O'Brien JP: The Oswestry low back pain disability questionnaire. **Physiotherapy 66:**271–273, 1980
13. Geesink RG, Hoefnagels NH, Bulstra SK: Osteogenic activity of OP-1 bone morphogenetic protein (BMP-7) in a human fibular defect. **J Bone Joint Surg Br 81:**710–718, 1999
14. Grauer JN, Patel TC, Erulkar JS, Troiano NW, Panjabi MM, Friedlaender GE: 2000 Young Investigator Research Award winner. Evaluation of OP-1 as a graft substitute for intertransverse process lumbar fusion. **Spine 26:**127–133, 2001
15. Grauer JN, Vaccaro AR, Kato M, Kwon BK, Beiner JM, Patel TC, et al: Development of a New Zealand white rabbit model of spinal pseudarthrosis repair and evaluation of the potential role of OP-1 to overcome pseudarthrosis. **Spine 29:**1405–1412, 2004
16. Hägg O, Fritzell P, Nordwall A: The clinical importance of changes in outcome scores after treatment for chronic low back pain. **Eur Spine J 12:**12–20, 2003
17. Hak DJ, Makino T, Niikura T, Hazelwood SJ, Curtiss S, Reddi AH: Recombinant human BMP-7 effectively prevents non-union in both young and old rats. **J Orthop Res 24:**11–20, 2006
18. 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 27:** 2654–2661, 2002
19. Kanayama M, Hashimoto T, Shigenobu K, Yamane S, Bauer TW, Togawa D: A prospective randomized study of posterolateral lumbar fusion using osteogenic protein-1 (OP-1) versus local autograft with ceramic bone substitute: emphasis of surgical exploration and histologic assessment. **Spine 31:**1067–1074, 2006
20. Li T, Surendran K, Zawaideh MA, Mathew S, Hruska KA: Bone morphogenetic protein 7: a novel treatment for chronic renal and bone disease. **Curr Opin Nephrol Hypertens 13:**417–422, 2004
21. Magin MN, Delling G: Improved lumbar vertebral interbody fusion using rhOP-1: a comparison of autogenous bone graft, bovine hydroxylapatite (Bio-Oss), and BMP-7 (rhOP-1) in sheep. **Spine 26:**469–478, 2001
22. Maric I, Poljak L, Zoricic S, Bobinac D, Bosukonda D, Sampath KT, et al: Bone morphogenetic protein-7 reduces the severity of colon tissue damage and accelerates the healing of inflammatory bowel disease in rats. **J Cell Physiol 196:**258–264, 2003
23. Moazzaz P, Gupta MC, Gilotra MM, Gilotra MN, Maitra S, Theerajunyaporn T, et al: Estrogen-dependent actions of bone morphogenetic protein-7 on spine fusion in rats. **Spine 30:** 1706–1711, 2005
24. Mofidi A, Sedhom M, O'Shea K, Moore D, Fogarty E, Dowling F: Usefulness of functional outcome questionnaires in primary screening of spinal disease. **J Spinal Disord Tech 16:**66–70, 2003
25. Patel TC, Erulkar JS, Grauer JN, Troiano NW, Panjabi MM, Friedlaender GE: Osteogenic protein-1 overcomes the inhibitory effect of nicotine on posterolateral lumbar fusion. **Spine 26:** 1656–1661, 2001
26. Patrick DL, Deyo RA, Atlas SJ, Singer DE, Chapin A, Keller RB: Assessing health-related quality of life in patients with sciatica. **Spine 20:**1899–1909, 1995
27. Rampersaud YR, Moro ER, Neary MA, White K, Lewis SJ, Massicotte EM, et al: Intraoperative adverse events and related postoperative complications in spine surgery: implications for enhancing patient safety founded on evidence-based protocols. **Spine 31:**1503–1510, 2006
28. Reddi AH: Role of morphogenetic proteins in skeletal tissue engineering and regeneration. **Nat Biotechnol 16:**247–252, 1998
29. Riddle DL, Stratford PW: Use of generic versus region-specific functional status measures on patients with cervical spine disorders. **Phys Ther 78:**951–963, 1998
30. Ripamonti U: Bone induction by recombinant human osteogenic protein-1 (hOP-1, BMP-7) in the primate Papio ursinus with expression of mRNA of gene products of the TGF-beta superfamily. **J Cell Mol Med 9:**911–928, 2005
31. Slover J, Abdu WA, Hanscom B, Weinstein JN: The impact of comorbidities on the change in short-form 36 and oswestry scores following lumbar spine surgery. **Spine 31:**1974–1980, 2006
32. Urist MR: Bone: formation by autoinduction. **Science 150:** 893–899, 1965
33. Vaccaro AR, Anderson DG, Patel T, Fischgrund J, Truumees E, Herkowitz HN, et al: Comparison of OP-1 Putty (rhBMP-7) to iliac crest autograft for posterolateral lumbar arthrodesis: a minimum 2-year follow-up pilot study. **Spine 30:**2709–2716, 2005
34. Vaccaro AR, Anderson DG, Toth CA: Recombinant human osteogenic protein-1 (bone morphogenetic protein-7) as an osteoinductive agent in spinal fusion. **Spine 27 (1 Suppl):**S59–S65, 2002
35. Vaccaro AR, Patel T, Fischgrund J, Anderson DG, Truumees E,

Herkowitz H, 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 12:**495–500, 2003

36. Vaccaro AR, Patel T, Fischgrund J, Anderson DG, Truumees E, Herkowitz HN, et al: A pilot study evaluating the safety and efficacy of OP-1 Putty (rhBMP-7) as a replacement for iliac crest autograft in posterolateral lumbar arthrodesis for degenerative spondylolisthesis. **Spine 29:**1885–1892, 2004
37. Walsh TL, Hanscom B, Lurie JD, Weinstein JN: Is a condition-specific instrument for patients with low back pain/leg symptoms really necessary? The responsiveness of the Oswestry Disability Index, MODEMS, and the SF-36. **Spine 28:**607–615, 2003
38. Ware JE Jr: Conceptualizing and measuring generic health outcomes. **Cancer 67 (3 Suppl):**774–779, 1991
39. Ware JE Jr, Kosinski M: **SF-36 Physical and Mental Health Summary Scales: A Manual for Users of Version 1, ed 2.** Lincoln, RI: QualityMetric Incorporated, 2001
40. Yang S, Zhong C, Frenkel B, Reddi AH, Roy-Burman P: Diverse biological effect and Smad signaling of bone morphogenetic protein 7 in prostate tumor cells. **Cancer Res 65:**5769–5777, 2005

Manuscript submitted May 9, 2007.
Accepted July 17, 2007.
This work was supported by funds from the Lawson Fellow in Neurology from The Toronto General & Western Hospital Foundation (J.C.F.), the Henry A. Beatty Scholarship (J.C.F.), the Krembil Chair in Neural Repair and Regeneration (M.G.F.), and Stryker Biotech (M.G.F.).

Address correspondence to: Michael G. Fehlings, M.D., Ph.D., Division of Neurosurgery, Department of Surgery, Toronto Western Hospital, University of Toronto, 399 Bathurst Street, 4W449, Toronto, Ontario M5T 2S8, Canada. email: Michael.Fehlings@uhn.on.ca.

FDA DEVICE/DRUG STATUS: INFUSE: This abstract does not discuss or include any applicable devices or drugs.

doi:10.1016/j.spinee.2008.06.096

80. Osteogenic Differentiation of Adipose Derived Stromal Cells Induction of Spinal Fusion in a Rodent Model

Jun Zou, MD, MS[1], Huilin Yang, MD, PhD[1], Dechun Geng, BS[1], Xuesong Zhu, MS[1], Ling Liu, MD, MS[1], Masashi Miyazaki, MD, PhD[2], Jeffrey Wang, MD, MSci[2]; [1]The First Affiliated Hospital of Soochow University, Suzhou, China; [2]University of California, Los Angeles, Santa Monica, CA, USA

BACKGROUND CONTEXT: Spinal arthrodesis is a fundamental treatment for many spinal pathologies. Stem cells from bone marrow have been used to stimulate bone formation for spine fusion However, harvesting bone marrow in human involved with potential donor site morbidity as well as harvesting autogenous bone graft. Recently, it has been demonstrated that adipose tissue contains multipotent stem cells and represents an alternative stem cell source to bone marrow mesenchymal stem cells. In addition to an adequate supply, the number of adipose derived stromal cells (ADSCs) does not appear to decline with age.
PURPOSE: To study the ability of adipose derived stromal cells to differentiate into osteogenic cells in vitro and induce spine fusion in a rat model.
STUDY DESIGN/ SETTING: This is an in vitro basic science study and an in vivo animal study.
OUTCOME MEASURES: In vitro study: Osteogenic differentiation was confirmed using the ALP and Von-Kossa Staining. Expressions of osteoblast specific genes (ALP, Osteopontin, and Osteocalcin) were confirmed by RT-PCR. In vivo study: At week 8, rat spinal fusion after implantation of these cells was assessed by X-ray, CT scan, manual palpation and biomechanical testing.
METHODS: In vitro study: adipose derived stromal cells were isolated from rat inguinal fat pads after extensive washing with phosphate derived saline and digesting with collagenase. After primary culture in osteogenic medium and expanded to two passages, the cells were incubated in either an osteogenic medium for 2–4 weeks to induce osteogenesis. In vivo study: Total 32 Sprague-Dawley male rats were underwent posterolateral lumbar fusions with implantation of materials in the intertransverse process space at L4-5. Group I: (n=8) rats were implanted with 1×107 ADSCs (P2)+collagen sponge and Group II: (n=8) rats were implanted with collagen sponge only. Group III, IV: (each group, n=8) rats were implanted with autograft or sham surgery with decortication of the transverse processes only.
RESULTS: In vitro study: ADSCs were able to be isolated from rat adipose tissue and expanded rapidly. It exhibited a heterogeneous population of fibroblast like cells morphologically. ADSCs induced to osteogenesis were stained positively for ALP activity after 2 weeks and formed mineralized nodular structures, as conformed by Von Kossa staining. Expression of osteoblast specific genes, such as ALP, Osteopontin, Osteocalcin, were detected. ALP and Osteopontin, were expressed constitutively in osteogenic medium after 2 and 4 weeks of culture. Expression of Osteocalcin, was induced by osteogenic growth factors at 4 weeks. In vivo study, All of the rats in Group I (ADSCs group) were judged to be completely fused by radiographic analysis (X-rays and CT) and manual palpation. Minimal or no evidences of bone formation were observed in Group II, III, IV. With biomechanical testing (Extension, Flexion, Lateral Bending), greater stiffness was shown in Group I compared to Groups II, III, IV. ($p<0.05$).
CONCLUSIONS: ADSCs can be isolated from rat adipose tissue. Their biological characteristics are similar with bone marrow mesenchymal stem cells, and have the potential to differentiate into osteogenic lineage both in vitro and in vivo. These cells also have the ability to induce a spinal fusion in a rat intertransverse process fusion model. This may prove to be an attractive strategy for bone formation and spinal fusion in humans. Further study is needed.

FDA DEVICE/DRUG STATUS: This abstract does not discuss or include any applicable devices or drugs.

doi:10.1016/j.spinee.2008.06.097

81. Effect of Teriparatide rhPTH134 and Calcitonin on Intertransverse Process Fusion in a Rabbit Model

Ronald Lehman, Jr., MD[1], Anton Dmitriev, MSc[2], Mario J. Cardoso, MD[2], Jolynn Raymond, DVM[3], Christen Christensen, DVM[3], Melvin Helgeson, MD[1], Timothy Kuklo, MD[4], K. Daniel Riew, MD[5]; [1]Walter Reed Army Medical Center, Washington, DC, USA; [2]Washington, DC, USA; [3]Armed Forces Institute of Pathology, Washington, DC, USA; [4]St. Louis, MO, USA; [5]Washington University in St. Louis, St. Louis, MO, USA

BACKGROUND CONTEXT: It is widely recognized that some osteoporosis medications, including bisphosphonates, can interfere with bone healing. Although prescribed frequently in the treatment of osteoporosis, the effect of teriparatide and calcitonin on spinal fusion has not been fully elucidated.
PURPOSE: To evaluate the effect of teriparatide and calcitonin following an intertransverse process spinal fusion in a rabbit model.
STUDY DESIGN/ SETTING: Randomized, double-blinded, placebo controlled animal study.
PATIENT SAMPLE: 51 New Zealand White Rabbits.
OUTCOME MEASURES: Emery histological scale, manual palpation, radiographic evaluation, and biomechanical evaluation.
METHODS: Fifty-one New Zealand white (NZW) rabbits underwent a posterolateral L5-L6 intertransverse process arthrodesis using autogenous iliac crest bone graft. The rabbits were randomly divided into three groups. All animals received daily subcutaneous injections of: Group I (n=17) 1cc of saline placebo; Group II (n=17) 10mcg/kg/d of teriparatide; Group III (n=17) 14IU/animal of calcitonin during the 8-week postoperative period. Post-mortem analyses included manual palpation, radiographic, biomechanical and histologic assessment (performed by two independent veterinary pathologists). Three random 10x fields were examined/graded within the cephalad, middle and caudad regions of each section (828 fields). Fusion quality was graded using the Emery histological scale (0-7 based on fibrous/bone content of the fusion mass).
RESULTS: Histologic fusion rates for teriparatide averaged 86.7% and was significantly greater than the autograft control group (62.5%) ($p=0.033$). The average Emery grading score was 5.99+1.46(SD) for the autologous group and 6.26+0.93(SD) for the Forteo group ($p=0.031$). Radiographically, there was a strong trend towards teriparatide being superior to the calcitonin group (85.7% versus 56.3%, respectively; $p=0.07$). Although not significant, the Forteo group showed less motion in both flexion/extension (Group I: 8.8+1.3, Group II: 75+1.3 & Group III: 8.7 +1.3) and axial rotation (Group I: 2.3+1.1, Group II: 1.7+0.8, Group III: 2.1+0.7) ($p=0.118$).
CONCLUSIONS: Our results suggest that Forteo enhances spinal fusion while calcitonin has a neutral effect. The teriparatide group had the best histologic fusion rate and Emery scores, while the calcitonin group was similar to the saline controls. Although not significant, the Forteo group had a strong trend towards superior radiographic fusion over the calcitonin group.
FDA DEVICE/DRUG STATUS: Forteo: Approved for this indication; Calcitonin: Approved for this indication.

doi:10.1016/j.spinee.2008.06.098

82. Posterolateral Intertransverse Lumbar Fusion in Nonhuman Primates Using OP-1 Putty: Effect of Dose Concentration on Fusion Rate

Louis Jenis, MD[1], Dean Falb, PhD[2], Bryan Cunningham, MSc[3], Allen Pierce[2], Mary E.P. Goad, DVM[2], Denis Schrier, PhD[2]; [1]Boston Spine Group, Boston, MA, USA; [2]Hopkinton, MA, USA; [3]Towson, MD, USA

BACKGROUND CONTEXT: Autograft has been considered the ''gold standard'' for posterolateral intertransverse process fusion (PLF) although arthrodesis success is inconsistent. Effective alternatives to autograft would eliminate the need for a second surgical site and the associated donor site pain and morbidity. Osteogenic Protein-1 (rhBMP-7) has been utilized as a bone graft substitute for autograft in spinal fusion in numerous higher and lower animal models and anatomic sites.
PURPOSE: The purpose of this study is to determine the effective dose concentration of OP-1 in a primate model of instrumented PLF as a means of translating to human indications.
STUDY DESIGN/ SETTING: In vivo primate lumbar fusion study. Animal Care Committee approval was obtained.
OUTCOME MEASURES: Plain radiographic and fine cut computed tomographic imaging was obtained at various time courses throughout the study period and fusion determined by qualitative and quantitative analysis. In addition, biomechanical assessments of fusion were performed.
METHODS: Twenty-four adult male baboons were surgically treated with exposure of L4L5, placement of autograft harvested from the iliac crest (4.5ml/side) or a constant volume of OP-1 putty (6ml/side), and bilateral pedicle screw-rod fixation. The OP-1 treated animals (4/group) were assigned as follows: carrier only (2g type I collagen– 460mg carboxymethylcellulose) or carrier with OP-1 including 0.33mg/ml, 1.0mg/ml, 2.0mg/ml, or 4.0mg/ml. Animals were monitored over 4 months and then euthanized.
RESULTS: The carrier only group did not achieve fusion at 3 or 4-month time points. There was no difference in autograft or OP-1 groups with 1.0, 2.0 or 4.0mg/ml BMP-7 treatments (100% clinical fusion rate and grade). The 0.33mg/ml treatment fusion rate was similar to higher OP-1 doses by 4 months (75% fusion rate). No statistical differences were found between BMP-7 doses and autograft for any mode of biomechanical testing. Quantified CT was utilized to determine bony fusion mass volume (cortical and trabecular bone), morphology of the fusion tissue, and connectivity between the cranial and caudal transverse processes. The results confirm the 1.0, 2.0 or 4.0mg/ml BMP-7 treatment groups had more than twice the volume of bone within the fusion mass compared to the autograft treatment group (ICBG). Similar results were found for both total trabecular bone volume and total fusion volume on the right and left side.
CONCLUSIONS: The importance of optimizing dose of BMP-induced spinal fusion and carrier / bulking agent has been shown in the current study. When a constant volume of carrier is introduced in PLF, it appears that the concentration of OP-1 is a critical determinant of osteoinduction and clinical fusion. These studies allow for extrapolation to human administration of OP-1 as a bone graft substitute in PLF. The optimal primate dosages are likely to be highly predictive of response to new bone formation in human clinical trials.
FDA DEVICE/DRUG STATUS: OP-1: Approved for this indication.

doi:10.1016/j.spinee.2008.06.099

Thursday, October 16, 2008
4:15–5:15 PM
Concurrent Session 2: Cervical

83. Comparison of Three Methods of ADCF Using Rigid Plates Dynamic Compression Plates and Cages

Kyung-Soo Suk, MD[1], Ki-Tack Kim[1], Jung-Hee Lee[1], Sang-Hun Lee, MD[1], Jin-Soo Kim[2], Chan-Wan Park[2]; [1]Department of Orthopaedic Surgery, Kyung Hee University College of Medicine, Seoul, South Korea; [2]Seoul, South Korea

BACKGROUND CONTEXT: ACDF using plate fixation has many advantages including high fusion rates and prevention of graft extrusion. Cervical plates are evolving since locking mechanism was introduced. Recent advance in cervical plates is axial loading by dynamization. Recently stand-alone cage is also introduced for ACDF. However, there is no study comparing 3 fusion methods.
PURPOSE: The purpose of this study was to compare the outcomes of the 3 methods of ACDF (rigid plate, dynamization plate, and cage) and find a useful method.
STUDY DESIGN/ SETTING: Prospective randomized study.
PATIENT SAMPLE: Consecutive 96 patients who were planned to undergo one level ACDF due to degenerative cervical disc disease were studied prospectively. The patients were randomized into 1 of 3 treatment groups: group 1 (rigid plate; n=31); group 2 (dynamization plate; n=29); and group 3 (cage; n=36).
OUTCOME MEASURES: Clinical and radiological outcome was measured preoperatively, postoperatively and at 2 year follow up. Clinical outcome was measured by subjective improvement rate (%), neck pain by VAS, arm pain by VAS, dysphasia, donor site pain, postop complication, and medications. Radiological outcome was measured by height of fusion segment, segmental angle of fusion segment, fusion, collapse of graft, collapse of endplate.
METHODS: Clinical and radiological outcomes were compared among 3 methods of ACDF group.
RESULTS: There are no significant differences in subjective improvement rate, neck pain, arm pain score, dysphasia, donor site pain, postoperative complications and postoperative medications. No patient complained donor site pain in group 3. 4 patients complained dysphasia in plate fixation group and only one patients complained dysphasia in cage group. Pseudarthrosis was found in only one patients of group 1. There was no Pseudarthrosis in group 2 and 3. Graft collapse was found in 3 patients of group 1 and 2 respectively. Bony endplate collapse was found in 6 patients of group 3. There were no significant differences in height of fusion segment, segmental angle of fusion segment among three groups.
CONCLUSIONS: No significant differences in clinical results and union rates were found among the 3 methods of ACDF. Cage group had less dysphasia and less donor site pain. But cage group had higher risk of endplate collapse.
FDA DEVICE/DRUG STATUS: This abstract does not discuss or include any applicable devices or drugs.

doi:10.1016/j.spinee.2008.06.101

84. Lower Incidence of Dysphagia with Cervical Arthroplasty Compared to ACDF in a Prospective Randomized Clinical Trial

Paul C. McAfee, MD[1], Andrew Cappuccino, MD[2], John DeVine, MD[3], John Regan, MD[4], Frank Phillips, MD[5]; [1]Towson, MD, USA; [2]Buffalo, NY, USA; [3]Tacoma, WA, USA; [4]Beverly Hills, CA, USA; [5]Chicago, IL, USA

BACKGROUND CONTEXT: This is a report of 132 patients from five investigational centers in the FDA prospective trial using a validated dysphagia outcomes instrument.
PURPOSE: The dysphagia data for both PCM and ACDF patients was reviewed from 5 centers to [1] compare the severity of dysphagia, [2] compare the postoperative incidence of dysphagia, and [3] to compare the resolution of perioperative dysphagia.
STUDY DESIGN/ SETTING: Level I Prospective Randomized Clinical trial.
PATIENT SAMPLE: This is a report of 132 patients from five investigational centers in the FDA prospective trial using a validated dysphagia outcomes instrument.
OUTCOME MEASURES: NDI,VAS, Bazaz Criteria.
METHODS: Patients between 18 and 65 years old with one-level symptomatic cervical radiculopathy and/or myelopathy for progressive neurological symptoms, were randomized to undergo anterior decompression and PCM arthroplasty or ACDF (control). Patients self-reported dysphagia severity using a Bazaz scale preoperatively and at follow-up. The Bazaz

SPINE Volume 20, Number 21, pp 2313–2317
©1995, Lippincott–Raven Publishers

Evaluation of Lumbar Spine Fusion

Plain Radiographs Versus Direct Surgical Exploration and Observation

Andrew P. Kant, MD,* Wayne J. Daum, MD,† S. Michael Dean, MD,* and Tatsuo Uchida, MS‡

Study Design. In a retrospective study, the incidence of false positive and false negative interpretation of x-rays for solid spinal arthrodesis with spinal instrumentation was evaluated in 75 patients.

Objective. To evaluate the accuracy of the interpretation of x-rays for diagnosing solid spinal arthrodesis in patients with spinal instrumentation.

Summary of Background Data. This retrospective study compared spinal fusion, as determined by direct observation and radiographic evaluation, in 75 patients with instrumented lumbar fusions using multiple devices. The fusions included posterolateral fusions or posterolateral with interbody fusions. Autograft, allograft, and a combination of these also were used.

Methods. A single blinded examiner reviewed all x-rays immediately before the spinal hardware was removed and the fusion mass was explored by the surgeon.

Results. There was a positive correlation between x-rays and the observations at the time of surgery in only 68% of the patients.

Conclusion. This study indicates that the accuracy of x-ray interpretation for spinal arthrodesis is only 68%. The L4–L5 level was the most difficult level to fuse and the most difficult to interpret using x-rays. Patients with persistent back pain, when nonmechanical causes have been ruled out, should be considered for surgical exploration of the fusion mass even if x-rays appear to indicate a solid fusion. [Key words: radiograph, spinal arthrodesis, spinal instrumentation] **Spine 1995;20: 2313–2317**

Producing a solid spinal fusion is often the goal of spinal surgical procedures. Radiographic fusion is not routinely associated with clinical success, or conversely, the radiographic failure of bony fusion is not uniformly associated with clinical failure. Postoperative failure, manifested by persistent or recurrent pain, may occur in 30–40% of patients.[9,14] Pseudarthrosis is thought to be a cause of a significant number of these clinical failures.[5,9,11,14] Determining the presence of arthrodesis at all spinal levels approached has been problematic (Figures 1A, B). It is even more difficult to evaluate the occurrence of solid spinal fusion in the presence of spinal instrumentation (Figures 2A, B).

Various imaging techniques have been evaluated. Plain films, accompanied by flexion and extension bending views, are commonly used.[14–16] Digitized or stereophotogrammetric analysis of vertebral motion segments,[13,18] plain tomography,[4] and computerized tomography with its various permutations[3,11,12] also have been examined. Plain and SPECT scintigraphy have been used to examine for pseudarthrosis.[17] However, to the best of our knowledge, only two studies have compared imaging modalities with the most objective standard of direct observation and manipulation of the spinal fusion mass.[1,2] The present retrospective study was undertaken to evaluate the accuracy of x-ray interpretation compared with intraoperative observation of the fusion mass in the presence of spinal instrumentation.

The senior author (APK) removed hardware at approximately 1 year after fusion for all of the patients reported in this study. This gave us the opportunity to compare the accuracy of plain radiograph and direct observation interpretation in a series of patients who had undergone fusion with instrumentation.

Methods

Over 300 lumbar spine arthrodesis procedures, using spinal instrumentation, have been performed by the senior author. The patients reviewed in this study include 75 who had hardware removal between 1 month and 212 weeks. One patient, who had hardware removal at 1 month, was eliminated from the study and is not included in the figures. Five other patients had hardware removed and the fusion explored before 10 months.

We want to emphasize that the purpose of this retrospective study was to compare the accuracy of radiographic observation with that of direct intraoperative observation of the fusion mass, not to examine solely for the occurrence of fusion. In all cases, the patients had some persistent back pain. The median time of hardware removal was 52 weeks, with an average of 51 weeks. The principal spinal instrumentation system used was the VSP Steffee instrumentation system. The Luque instrumentation system was used in four patients. For

From *KSF Orthopaedic Center, Houston, Texas, †the Department of Orthopaedics, University of Texas Houston Health Science Center, Houston, and ‡the Office of Biostatistics, University of Texas Medical Branch, Galveston, Texas.
Accepted for publication April 25, 1995.
Device status category: 1.

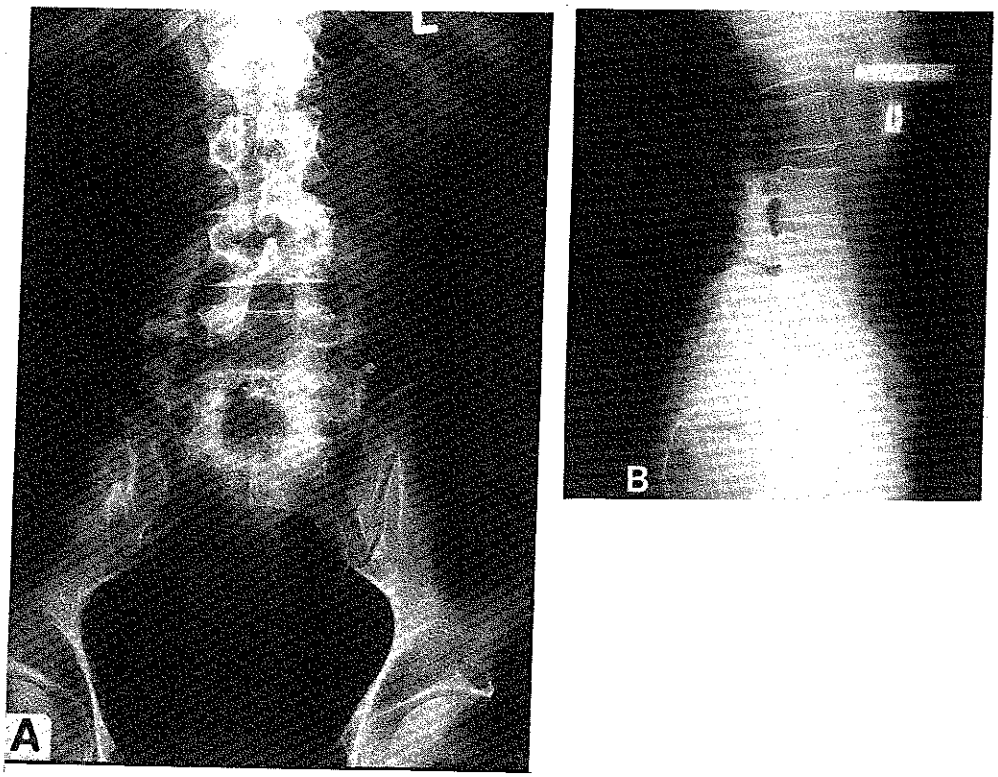


Figure 1. (**A**) Anteroposterior view of lumbar spine. Obvious pseudarthrosis at L4–L5. Apparent solid fusion at L3–L4. However, direct observation at surgery indicated pseudarthrosis at L4–L5 and at L3–L4. (**B**) Lateral view of lumbar spine. Spondylolisthesis at L4–L5. Occurrence of fusion is difficult to determine.

three thoracolumbar fractures, Moe Harrington Rod systems were used.

In all cases, posterolateral fusion was performed. This included complete exposure of the transverse process, lateral wall of the facet, and sacral ala where appropriate. All of these surfaces were decorticated. In addition, osteotomy of the facet joints, with removal of articular cartilage, was performed at the level to be fused. In 37 patients, posterior interbody fusion also was performed. The posterior lumbar interbody fusion in these patients was performed with cancellous bone chips. In an additional six patients, posterior lumbar interbody fusion with allograft was performed as a sole procedure without postero-

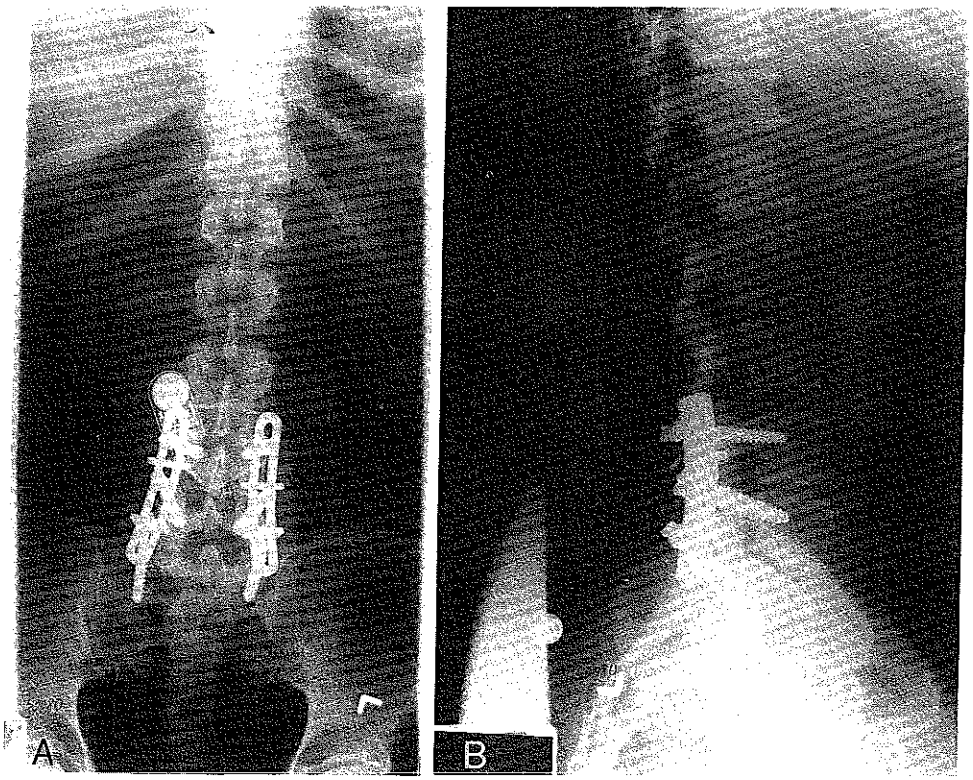


Figure 2. (**A**) Anteroposterior view of lumbar spine. Apparent solid fusion at L4–L5 and L5–S1. At the time of hardware removal and surgical exploration, a pseudarthrosis at L4–L5 was found. (**B**) Lateral view of lumbar spine. Posterior lumbar interbody fusion was attempted at L4–L5. Interbody fusion is not solid.

lateral bone. Besides the six patients in whom posterior lumbar interbody fusion was used, allograft was used in the posterolateral position in an additional two patients. No anterior lumbar interbody fusions were performed in this study.

All patients were given the option of having the hardware removed. The hardware generally was not removed until at least 10 months from the time of surgery. However, in one patient, the hardware was removed at 1 month because of persistent radicular symptoms believed to be related to screw placement. Obviously, fusion would not be solid in 1 month. Therefore, this patient was eliminated from the study. Hardware removal was suggested for those patients who had persistent back pain, persistent radicular symptoms, suggestion of pseudarthrosis on radiographs, or infection. There was one case of infection, necessitating hardware removal. However, the plate was not removed until 1 year after surgical intervention. Therefore, this patient remained in this study.

Direct surgical exploration and observation of the fusion mass was performed on 126 vertebral levels in 75 patients at the time of hardware removal. Direct surgical exploration of the fusion mass was carried out by removing the hardware and removing all soft tissue over the fusion mass. An uninvolved orthopedic surgeon viewed the last set of plain radiographs (five views) before the hardware removal procedure in these patients. The last set of radiographs was performed between 1 and 4 weeks before hardware removal. In most cases, an anteroposterior Ferguson view was included.

The observer's impression of the solidity of the fusion at each individual level was recorded. Radiographic determination of fusion is mostly subjective. However, a fusion was considered solid radiographically only when solid bone could be visualized from one transverse process to the other transverse process or when oblique views revealed obliteration and fusion of the facet joint. This interpretation was compared with the results of direct surgical observation of the solidity of the fusion mass at each level in every patient.

At the time of surgery, solid fusion was indicated only if both of the following criteria were met.

1. No motion. Solidity of the fusion was determined by several different intraoperative methods.
   A. A towel clip was placed on the intact spinous process. Attempts were made to move the segment. Towel clips were used to distract and torque and compress the fusion mass. If motion was seen, a pseudarthrosis was diagnosed.
   B. If the canal was opened, an instrument was placed in the disc space. The instrument was manipulated to separate the vertebral segments. If there was motion from one vertebral segment to another, a pseudarthrosis was diagnosed.
   C. Direct pressure was placed on the fusion mass and facet with a large punch. If motion was seen, a pseudarthrosis was diagnosed.
   D. Large 10-inch long punches were placed in each screw hole. Distraction, compression, and torque were applied along these large lever arms. Any motion indicated a pseudarthrosis.

If motion was seen from any of these four methods, careful exploration of the fusion mass always revealed the presence of a synovial or fibrous pseudarthrosis.

Table 1. Summary of Direct Observation

Level	Procedures Performed (No.)	Not Fused		Fused	
		No.	Percent	No.	Percent
Above L3	8	3	37.5	5	62.5
L3–L4	11	4	36.4	7	63.6
L4–L5	61	23	37.7	38	62.3
L5–S1	46	9	19.6	37	80.4*
Total	126	39	30.9	87	69.1

* The rate of fusion at this level is statistically higher than that at other levels ($P < 0.05$), according to the chi-squared test.

2. Direct surgical exploration and observation of the fusion mass. The fusion mass was thoroughly explored. All soft tissue was removed from the area of the fusion mass. The facet joint was explored. The intertransverse process area also was explored. For an arthrodesis to be determined as solid, a solid mass of bone without interruption had to be visualized from ala to transverse process, or there had to be total obliteration of the facet joint with overgrowth of bone to both sides.

Statistical Methods

Fusion rates at four different vertebral levels (L5–S1, L4–L5, L3–L4, and those above L3) were compared using the chi-squared test and a method for partitioning total chi-squared results into subgroups for isolating sources of significant differences.[6]

Kappa statistics were computed to measure agreement between two raters or two rating methods (direct observation versus radiographic interpretation) of fusion. The magnitude of the Kappa statistic was interpreted as follows.[7]

$0.75 < \text{Kappa} < 1.00$: excellent agreement beyond chance.

$0.40 < \text{Kappa} < 0.75$: fair to good agreement beyond chance.

$\text{Kappa} < 0.40$: poor agreement.

Results

The results of our study of 126 levels in 75 patients are summarized in Tables 1 and 2. According to direct observation, 87 of 126 levels were fused, representing an overall fusion rate of 69%. Thirty-nine levels were not fused at the time of direct observation. There were 46 attempted fusions at L5–S1, 61 at L4–L5, 11 at L3–L4, and eight above the L3–L4 level. The numbers of actual fusions, which were confirmed by direct observation at the time of hardware removal, were 37 at L5–S1, for a fusion rate of 80%; 38 at L4–L5, for a fusion rate of 62%; seven at L3–L4, for a 64% fusion rate; and five above the L3–L4 level, for a 63% fusion rate. The fusion rate at L5–S1 was significantly greater than at the other levels (Table 1).

Agreement between direct observation and radiographic interpretation was fair at L3–L4 and L5–S1

Table 2. Summary of Agreement and Disagreement

		Agreement				Disagreement				
		DO, RI Not Fused		DO, RI, Fused		DO Not Fused, RI Fused		DO Fused, RI Not Fused		
Level	Total	No.	Percent	No.	Percent	No.	Percent	No.	Percent	Kappa Statistic
Above L3	8	1	12.5	4	50.0	2	25.0	1	12.5	0.14
L3–L4	11	2	18.2	7	63.6	2	18.2	0	0.0	0.56
L4–L5	61	6	9.8	29	47.5	17	27.9	9	14.8	0.03
L5–S1	46	6	13.0	34	73.9	3	6.5	3	6.5	0.59
Overall	126	15	11.9	74	58.7	24	19.1	13	10.3	0.26

DO = direct observation. RI = radiographic interpretation.

(Kappa statistic = 0.56 and 0.59, respectively). In contrast, the agreement was poor above L3 and very poor at L4–L5 (Kappa statistic = 0.14 and 0.02, respectively). Overall agreement was poor, indicated by a Kappa statistic of 0.26 with a 95% confidence interval of 0.07–0.44 (Table 2).

Discussion

The most significant finding of this study was the remarkably poor correlation between the interpretation of the plain radiographs and the actual fusion status of the spine, as determined by direct observation, regardless of the time interval from instrumentation to exploration. This study confirms the findings of Blumenthal and Gill[1] and Brodsky et al,[2] who reported series of comparable size to the one reported here. Of particular concern are those cases in which the radiographs were interpreted as indicating a solid fusion, but were really pseudarthrosis (Table 2). The same posterolateral procedure was carried out at each level, but all levels did not fuse with the same frequency. The most common level approached in the study was L4–L5. In our study, 61 procedures were performed at this level. Our study indicates that this level is the most difficult to fuse and the most difficult level to interpret using radiographs.

Studies are frequently presented that report clinical outcomes of spinal fusions based on radiographic evaluation. Many authors have reported series with high percentages of solid arthrodesis. However, some of these reports have indicated that patients with solid fusions do not have clinical improvement compared with patients who have pseudarthroses.[8,14,18] Our study suggests that these series may be flawed when they depend on radiographic analysis only to verify fusion.

There has been a significant increase in spinal surgery. Lumbar arthrodesis has become a much more frequently performed procedure during the last 10 years compared with the preceding decade. The procedure is expensive, not only in the actual dollars and cents of hospitalization, medications, physical therapy, and physician fees, but in time lost from work and activities by the patient. Several series have presented patient outcomes based on radiographic interpretation. Several of these series include the use of instrumentation and electrical stimulation.[9,10] The use of metal implants reduces the value of flexion and extension views. Other studies have not shown a substantially improved rate of pseudarthrosis detection using modalities such as tomography, computed tomography, and scintigraphy. No imaging study has been shown to be uniformly effective in determining pseudarthrosis.

This study demonstrates that using plain radiographs alone to determine the solidity of spinal fusion can be misleading. When a patient presents with persistent back pain (versus radicular pain) after a lumbar arthrodesis, pseudarthrosis may be the cause of that pain despite apparent radiographic evidence of fusion. As a result of these observations, we are more inclined to re-explore patients who have persistent back symptoms and whose results are clinical failures, even if the radiographic appearance of the fusion mass is solid in the presence of spinal instrumentation. This re-exploration is undertaken only after nonmechanical causes of pain are evaluated and ruled out.

References

1. Blumenthal SL, Gill K. Can lumbar spine radiographs accurately determine fusion in postoperative patients? Correlation of routine radiographs with a second surgical look at lumbar fusions. Spine 1993;18:1186–9.

2. Brodsky AE, Kovalsky ES, Khalil MA. Correlation of radiologic assessment of lumbar spine fusions with surgical exploration. Spine 1993;16S:261–5.

3. Chafetz N, Cann CE, Morris JM, Steinbach LS, Goldberg HI. Pseudarthrosis following lumbar fusion: detection by direct coronal CT scanning. Radiology 1987;162:803–5.

4. Dawson EG, Clader TJ, Bassett LW. A comparison of different methods used to diagnose pseudarthrosis following posterior spinal fusion for scoliosis. J Bone Joint Surg [Am] 1985;67:1153–9.

5. Davne SH, Myers DL. Results following removal or revision of lumbar pedicle screw instrumentation. Proceedings of the North American Spine Society Seventh Annual Meeting. Boston, Massachusetts, 1992:104–5.

6. Fleiss JL. Statistical Methods for Rates and Proportions. 2nd ed. New York: John Wiley & Sons, Inc., 1981:138–43.

7. Fleiss JL. Statistical Methods for Rates and Proportions. 2nd ed. New York: John Wiley & Sons, Inc., 1981:212–21.

8. Frymoyer JW, Hanley EN Jr, Howe J, Kuhlmann D, Matteri RE. A comparison of radiographic findings in fusion and nonfusion patients ten or more years following lumbar disc surgery. Spine 1979;4:435–40.

9. Grubb SA, Lipscomb HJ. Results of lumbosacral fusion for degenerative disc disease with and without instrumentation. Spine 1992;17:349–55.

10. Kane WJ. Direct current electrical bone growth stimulation for spinal fusion. Spine 1988;13:363–5.

11. Laasonen EM, Soini J. Low back pain after lumbar fusion surgical and computed tomographic analysis. Spine 1989;14: 210–3.

12. Lang P, Genant HK, Chafetz N, Steiger P, Morris JM. Three-dimensional computed tomography and multiplanar reformations in the assessment of pseudarthrosis in posterior lumbar fusion patients. Spine 1989;13:69–75.

13. Morris JM, Chafetz N, Baumrind S, Genant H, Korn EL. Stereophotogrammetry of the lumbar spine. Spine 1985;10: 368–75.

14. O'Beirne JD, O'Neill JG, Williams DH. Spinal fusion for back pain: a clinical and radiological review. J Spinal Disord 1992;5:32–8.

15. Pearcy M, Burrough S. Assessment of bony union after interbody fusion of the lumbar spine using a biplanar radiographic technique. J Bone Joint Surg [Br] 1982;64:228–32.

16. Ralston EL, Thompson WAL. The diagnosis and repair of pseudarthrosis of the spine. Surg Gynecol Obstet 1949:37–42.

17. Slizofski WJ, Collier BD, Flatley TJ, Carrera GF, Hellman RS, Isitman AT. Painful pseudarthrosis following lumbar spinal fusion: detection by combined SPECT and planar bone scintigraphy. Skeletal Radiol 1987:136–41.

18. Suzuki T, Pearcy MJ, Tibrewal SB, Wilson D, Duthie RB. Posterior intertransverse fusion assessed clinically and with biplanar radiography. Int Orthop 1985;9:11–17.

Address reprint requests to

Andrew P. Kant, MD
KSF Orthopaedic Center
17270 Red Oak Drive
Suite 200
Houston, TX 77090

SPINE Volume 29, Number 7, pp 726–734
©2004, Lippincott Williams & Wilkins, Inc.

Degenerative Lumbar Spondylolisthesis With Spinal Stenosis

A Prospective Long-Term Study Comparing Fusion and Pseudarthrosis

Martin B. Kornblum, MD,* Jeffrey S. Fischgrund, MD,† Harry N. Herkowitz, MD,† David A. Abraham, MD,‡ David L. Berkower, DO,§ and Jeff S. Ditkoff‖

Study Design. A prospective, randomized study on patients who underwent posterior lumbar decompression with bilateral posterolateral arthrodesis.

Objective. To determine the long-term influence of pseudarthrosis on the clinical outcome of patients with degenerative spondylolisthesis and spinal stenosis.

Summary of Background Data. Spinal decompression and posterolateral arthrodesis have been shown to be beneficial in the surgical treatment of symptomatic spinal stenosis with concurrent spondylolisthesis.

Methods. Forty-seven patients with single-level symptomatic spinal stenosis and spondylolisthesis were prospectively studied. Patients were treated with posterior decompression and bilateral posterolateral arthrodesis with autogenous bone graft. Radiographic evaluation was used to determine if fusion or pseudarthrosis was present. The solid fusion and pseudarthrosis groups were analyzed clinically, roentgenographically, and with a validated self-administered spinal stenosis questionnaire.

Results. Forty-seven patients were available for review at a range of follow-up from 5 to 14 years. Average follow-up was 7 years 8 months. Clinical outcome was excellent to good in 86% of patients with a solid arthrodesis and in 56% of patients with a pseudarthrosis ($P = 0.01$). Significant differences in residual back and lower limb pain was discovered between the two groups using a scale ranging from 0 (no pain) to 5 (severe pain). Preoperative back and lower limb pain scores were statistically similar between the two groups. The solid fusion group performed significantly better in the symptom severity and physical function categories on the self-administered questionnaire. The two groups had similar results in the patient satisfaction category of this questionnaire.

Conclusions. In patients undergoing single-level decompression and posterolateral arthrodesis for spinal stenosis and concurrent spondylolisthesis, a solid fusion improves long-term clinical results. Benefits of a successful arthrodesis over pseudarthrosis were demonstrated with respect to back and lower limb symptomatology compared with prior shorter-term studies, which indicated no significant difference in clinical outcome between the two groups. [Key words: degenerative spondylolisthesis, lumbar spine, spinal fusion, stenosis] **Spine 2004;29:726–734**

Degenerative spondylolisthesis was first described by Newman in 1955.[1] Earlier descriptions contrasted this condition from those caused by a pars interarticularis defect. Junghanns introduced the term "pseudospondylolisthesis" in 1930.[2] He recognized the distinction of an intact posterior element in his examination of anatomic specimens from Schmorl's collection. However, this term led to some confusion, as there is indeed a true spondylolisthesis in this condition. Thus, MacNab, in 1950, utilized "spondylolisthesis with an intact neural arch."[3] Wiltse *et al* established a widely accepted classification of spondylolisthesis based on etiology.[4] Degenerative spondylolisthesis comprises one of five elements in this system.

The operative management of degenerative spondylolisthesis has remained controversial. Early authors recommended decompression alone; stabilization procedures after laminectomy were considered unnecessary.[5–10] Herkowitz and Kurz, in 1991, performed a prospective, randomized study comparing decompression alone with decompression and bilateral posterolateral arthrodesis.[11] Fifty consecutive patients were assigned alternately to one of two treatment groups. Follow-up averaged 3 years. The results of this study demonstrated a significantly improved clinical outcome in those patients who underwent decompression with a concomitant arthrodesis. Pseudarthrosis was noted in 9 patients (36%) of the arthrodesis group. However, all patients with a pseudarthrosis had an excellent or good outcome at final evaluation.

The addition of spinal instrumentation has been advocated by some authors in the operative management of degenerative spondylolisthesis with spinal stenosis.[12–16] Instrumentation has been recommended to increase the fusion rate, decrease the rehabilitation time, and improve patient outcome.[17] However, based on the results of short to intermediate range studies, fusion status does not affect clinical outcome.[14,18] A fibrous union appears to provide sufficient stabilization and to provide pain relief of the back and lower extremities.

From the *Department of Orthopaedics, University of Virginia, Charlottesville, VA; †Southfield, MI; ‡West Reading, PA; §Department of Internal Medicine, Botsford Hospital, Farmington Hills, MI; and ‖Bloomfield Hills, MI.
Acknowledgment date: May 25, 2001. First revision date: November 12, 2001. Acceptance date: June 25, 2003.
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 Jeffrey S. Fischgrund, MD, 27207 Lasher Rd., Suite 200B, Southfield, MI 48034; E-mail: jsfischgrund@comcast.net

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

Fischgrund *et al*, in 1997, published a prospective, randomized study comparing the results of decompression and arthrodesis alone with those of decompression and arthrodesis combined with instrumentation.[18] Sixty-eight patients were randomized to one of two treatment groups. There was an average follow-up of 2 years. The results of this study demonstrated that the addition of spinal instrumentation will improve the fusion rate (83% *vs.* 45%). However, no significant improvement in clinical outcome was realized with the use of spinal instrumentation at final follow-up. Although pseudarthrosis developed in 55% of the noninstrumented group, the clinical outcome was still noted to be excellent or good in 15 of 18 patients (83%).[18]

The purpose of the current study was to determine the long-term influence of arthrodesis or pseudarthrosis on the clinical outcome of patients with degenerative spondylolisthesis and spinal stenosis. A prospective, randomized study was performed on patients who underwent posterior lumbar decompression with bilateral posterolateral autogenous arthrodesis.

■ Materials and Methods

A total of 118 consecutive patients had been randomly assigned to one of two treatment groups; these patients were described in two previous studies.[11,18] Fifty-eight patients from the prior two studies had been randomized to the treatment group that underwent posterior lumbar decompression and bilateral posterolateral autogenous arthrodesis without spinal instrumentation. The data on these 58 patients form the basis of this report.

All patients had degenerative spondylolisthesis with symptomatic spinal stenosis at a single level, with no prior history of lumbar spine surgery. All patients underwent a trial of nonoperative treatment for at least 3 months before surgery. The patients were recommended for a surgical procedure after failing nonoperative treatment. All continued to have significant back and leg pain with a significant restriction of daily activities due to radicular or neurogenic claudicatory complaints. Informed written consent was obtained from each participant. All patients in the current study agreed to participate in a clinical study approved by the Human Investigational Committee at William Beaumont Hospital, Royal Oak, MI.

Forty-seven of 58 patients treated with decompression and arthrodesis were available at final review. Final evaluation consisted of a telephone interview and self-administered questionnaire. Final clinical and radiographic assessment was performed approximately 3 years following surgery. Of the 11 patients not included in this report, 8 patients died, 1 had a recent cerebrovascular accident, 1 patient declined to participate, and only 1 patient was not located.

All patients had single-level degenerative lumbar spondylolisthesis on plain radiographic imaging. The diagnosis of spinal stenosis was established by computed tomography (CT), CT myelogram, or MRI. Preoperative plain radiographs of the lumbosacral spine were obtained for all patients. These included anteroposterior, lateral, left and right obliques, standing lateral, and standing flexion–extension lateral images. Final radiographs obtained included anteroposterior and standing flexion–extension lateral images. These final radiographs were obtained 3 years following surgical intervention. Preoperative and postoperative radiographic images were analyzed to determine the amount of spondylolisthesis in millimeters, the amount of sagittal motion in millimeters, and the amount of angular motion in degrees (Figure 1).

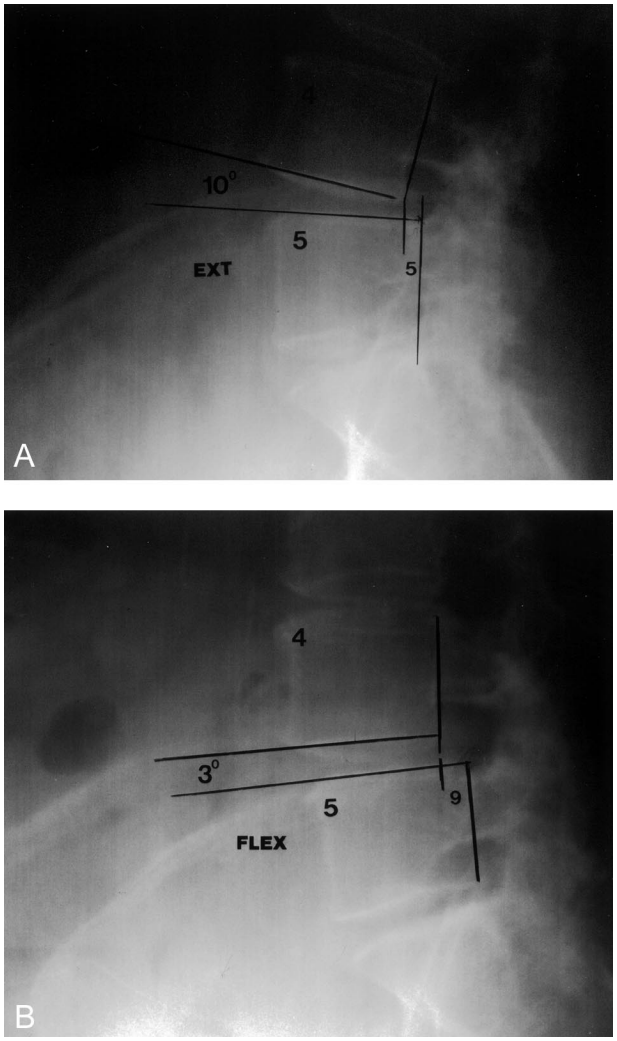


Figure 1. **A:** Preoperative lateral extension radiograph demonstrates 5 mm of subluxation of L4 on L5. **B:** The same patient demonstrates 9 mm of subluxation with flexion. This patient's sagittal motion would be 4 mm. The angular motion would be 7°.

Arthrodesis was determined to be successful if follow-up radiographs demonstrated a bilateral continuity in the fusion mass between the cephalad and caudad transverse processes. Pseudarthrosis was present if there was no continuity in the fusion mass (Figure 2) or if lateral flexion–extension radiographs demonstrated >2° of angular motion or >2 mm of sagittal motion at the level of the spondylolisthesis.[18] All clinical and radiographic assessments were made by examiners other than the treating surgeons and who were blinded to the patient's clinical results. Radiographs were independently examined by two orthopedic surgeons (one of whom was a spine specialist). If the reported fusion status differed between the examiners, the radiographs were reexamined and a consensus reached.

Decompression of the central canal and nerve roots was performed by removing half of the cephalad and the caudad lamina of the involved vertebra, together with bilateral medial caudad and cephalad facetectomy. The technique of spinal arthrodesis was that described by MacNab and Dall[19] and by

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

Figure 2. Two-year postoperative anteroposterior radiograph demonstrating clefts (arrows) in the lateral fusion mass between L4 and L5.

Table 1. Clinical and Radiographic Data

	Solid Fusion (N = 22)	Pseudarthrosis (N = 25)
Result		
Excellent	12 (54%)	7 (28%)
Good	7 (32%)	7 (28%)
Fair	2 (9%)	3 (12%)
Poor	1 (5%)	8 (32%)
Back pain		
Preop	3.7 (0–5)	3.5 (0–5)
Postop	1.4 (0–4)	2.6 (0–5)
Leg pain		
Preop	4.5 (3–5)	4.2 (0–5)
Postop	0.5 (0–3)	2.1 (0–5)
Olisthesis (mm)		
Preop	6.4 (2–18)	6.9 (2–15)
Postop	6.4 (2–14)	7.3 (2–15)
Sagittal motion (mm)		
Preop	3.2 (0–8)	3.3 (0–8)
Postop	1.0 (0–6)	2.6 (0–6)
Angular motion (°)		
Preop	6.6 (0–16)	10.1 (4–17)
Postop	0.5 (0–12)	8.4 (4–17)

Wiltse *et al*[20] for single-level bilateral intertransverse process arthrodesis. The outer table of the iliac crest was exposed through the same skin incision that was used for the decompression and arthrodesis. Strips of corticocancellous and cancellous bone were harvested from the outer and middle tables of the iliac crest and were placed along the transverse processes.[16,21] Decortication of the transverse processes with a burr or rongeur was performed before placement of bone graft.[18]

Before the operation, all patients rated pain in the back and lower limbs/buttocks on a visual analog pain scale, ranging from 0 (no pain) to 5 (severe pain). Separate scales were established for back and lower limb pain. At final follow-up, the patients were again asked to score their back and leg pain on the same visual analog pain scale.

The operative results were rated as excellent, good, fair, or poor based on criteria established from previous studies.[11,18] The result was considered to be excellent if a patient resumed unrestricted activity and had near-complete relief of pain in the back, lower limbs, or both. A good result indicated that there was occasional discomfort in the back or lower limbs, necessitating occasional non-narcotic medication. Patients with a good result had significant improvement, compared with the preoperative condition, and had resumed unrestricted activity. A fair result was defined as intermittent discomfort in the back, lower limbs, or both; improvement compared with the preoperative condition; restriction of activities; and an occasional need for non-narcotic medication. The patients who had a poor result had marked discomfort in the back, lower limbs, or both, necessitating non-narcotic and occasional narcotic medication. The patients in this category noted no improvement compared with the preoperative condition and had significant restriction of activities[18] (Table 1).

The clinical results of the operation and radiographic findings were then subjected to statistical analysis. Categorical variables were analyzed using the two-tailed Fisher's exact test or Cochran-Mantel-Haenszel statistics based on table scores where appropriate. Continuous variables were analyzed using the two-tailed Student's *t* test or the paired test. Pearson correlation coefficient and the asymptotic error were calculated as needed.

The same postoperative treatment was used for all groups of patients. Walking was permitted on the first postoperative day and progressed at 4 to 6 weeks after surgery. Exercises on the stationary bike or water therapy began at 6 to 8 weeks, and exercises for flexion of the spine and strengthening of the abdominal muscles were added at 10 to 12 weeks. No brace or corset was used after surgery in either group.[18] A self-administered spinal stenosis questionnaire, as developed by Stucki *et al*, was used to compare long-term postoperative outcome between the two groups.[22] This questionnaire was shown to be reproducible, internally consistent, valid, and highly responsive. Three categories are assessed *via* the questionnaire: symptom severity, physical function status, and patient satisfaction (Table 2).

Pain scales, operative results, and the self-administered spinal stenosis questionnaire were completed by the patients in a return trip to the hospital. Those patients who were unable to return to the hospital were administered the questionnaire through a telephone interview. This was performed by one of two medical students, who were unaware of the patient's fusion status at the time of this conversation. The duration of follow-up ranged from 5 to 14 years (mean 7.7 years).

■ Results

There were 36 women and 11 men in this study. The average age at surgery was 73 years for the solid fusion group and 72 years for the pseudarthrosis group. Nine patients were smokers, 8 patients were diabetic, and 6 patients had a diagnosis of peripheral vascular disease. The operations were performed at L4–L5 in 40 patients and at L3–L4 in 7 patients. The arthrodesis was successful in 22 patients (47%). Pseudarthrosis developed in 25 patients. Arthrodesis status was determined by radiographs taken at final clinical follow-up, usually 3 years after the surgical procedure (range 2–4 years).

Clinical outcome, assessed according to relief of pain and an increase in activity, was good or excellent in 86% of patients with a solid fusion and in 56% of patients

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

Table 2. Spinal Stenosis Questionnaire[22]

	Solid Fusion [no. (%)]	Pseudarthrosis [no. (%)]
I. **Symptom Severity Scale**		
In the last month, how would you describe:		
The pain you have had on average including pain in your back, buttocks and pain that goes down the legs?*		
None	8 (36)	2 (8)
Mild	9 (41)	9 (36)
Moderate	3 (14)	6 (24)
Severe	2 (9)	8 (32)
Very severe	0 (0)	0 (0)
How often have you had back buttock or leg pain?*		
Less than once a week	13 (62)	7 (28)
At least once a week	0 (0)	2 (8)
Every day, for at least a few minutes	6 (6)	4 (16)
Every day, for most of the day	2 (2)	11 (44)
Every minute of the day	0 (0)	1 (4)
The pain in your back or buttocks?*		
None	8 (36)	3 (12)
Mild	7 (22)	6 (24)
Moderate	5 (23)	8 (32)
Severe	1 (4)	8 (32)
Very severe	1 (5)	0 (0)
The pain in your legs or feet?*		
None	14 (64)	6 (24)
Mild	7 (32)	6 (24)
Moderate	1 (4)	7 (28)
Severe	0 (0)	6 (24)
Very severe	0 (0)	0 (0)
Numbness or tingling in your legs or feet?		
None	14 (64)	11 (44)
Mild	3 (14)	6 (24)
Moderate	4 (18)	5 (20)
Severe	1 (4)	3 (12)
Very severe	0 (0)	0 (0)
Weakness in your legs or feet?*		
None	13 (62)	8 (33)
Mild	4 (19)	4 (17)
Moderate	3 (14)	9 (38)
Severe	1 (5)	3 (12)
Very severe	0 (0)	0 (0)
Problems with your balance?		
No, I've had no problems with balance	14 (64)	11 (44)
Yes, sometimes I feel my balance is off, or that I am not sure-footed	7 (32)	9 (36)
Yes, often I feel my balance is off, or that I am not sure-footed	1 (4)	5 (20)
II. **Physical Function Scale**		
In the last month, on a typical day:		
How far have you been able to walk?*		
Over 2 miles	10 (46)	2 (8)
Over 2 blocks, but less than 2 miles	8 (36)	8 (33)
Over 50 feet, but less than 2 blocks	4 (18)	9 (38)
Less than 50 feet	0 (0)	5 (21)
Have you taken walks outdoors or in malls for pleasure?*		
Yes, comfortably	17 (77)	6 (24)
Yes, but sometimes with pain	1 (4.5)	8 (32)
Yes, but always with pain	1 (4.5)	4 (16)
No	3 (14)	7 (28)
Have you been shopping for groceries or other items?*		
Yes, comfortably	15 (68)	7 (28)
Yes, but sometimes with pain	3 (14)	7 (28)
Yes, but always with pain	1 (4)	6 (24)
No	3 (14)	5 (20)
Have you walked around the different rooms in your house or apartment?*		
Yes, comfortably	16 (73)	10 (40)
Yes, but sometimes with pain	6 (27)	9 (36)
Yes, but always with pain	0 (0)	6 (24)
No	0 (0)	0 (0)
Have you walked from your bedroom to the bathroom?*		
Yes, comfortably	18 (82)	11 (44)
Yes, but sometimes with pain	4 (18)	8 (32)
Yes, but always with pain	0 (0)	6 (24)
No	0 (0)	0 (0)

(*Table continues*)

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

Table 2. Continued

	Solid Fusion [no. (%)]	Pseudarthrosis [no. (%)]
III. **Satisfaction Scale**		
How satisfied are you with:		
The overall results of your operation?		
Very satisfied	16 (73)	12 (50)
Somewhat satisfied	4 (18)	4 (17)
Somewhat dissatisfied	0 (0)	2 (8)
Very dissatisfied	2 (9)	6 (25)
Relief of pain following the operation?		
Very satisfied	14 (64)	12 (50)
Somewhat satisfied	5 (23)	3 (12)
Somewhat dissatisfied	1 (4)	5 (21)
Very dissatisfied	2 (9)	4 (17)
Your ability to walk following the operation?		
Very satisfied	14 (64)	13 (52)
Somewhat satisfied	5 (23)	5 (20)
Somewhat dissatisfied	1 (4)	3 (12)
Very dissatisfied	2 (9)	4 (16)
Your ability to do housework, yardwork or job following the operation?		
Very satisfied	13 (62)	10 (40)
Somewhat satisfied	3 (14)	5 (20)
Somewhat dissatisfied	3 (14)	2 (8)
Very dissatisfied	2 (10)	8 (32)
Your strength in the thighs legs and feet?		
Very satisfied	13 (59)	9 (36)
Somewhat satisfied	4 (18)	8 (32)
Somewhat dissatisfied	4 (18)	4 (16)
Very dissatisfied	1 (5)	4 (16)
Your balance or steadiness on your feet?		
Very satisfied	14 (64)	12 (48)
Somewhat satisfied	5 (23)	6 (24)
Somewhat dissatisfied	1 (4)	2 (8)
Very dissatisfied	2 (9)	5 (20)

* Statistically significant better scores in the solid fusion group.

with a pseudarthrosis. There was a statistically significant difference in outcome ($P = 0.01$).

Significant differences were also demonstrated at final follow-up between the two groups with respect to residual back and limb pain scores. Preoperative back and leg pain scores were similar between the two groups. The solid fusion group and pseudarthrosis group had initial average back pain scores of 3.7 and 3.5, respectively. Preoperative leg pain scores were 4.5 and 4.2, respectively.

At the most recent postoperative evaluation, back pain scores were 1.4 and 2.6 ($P = 0.02$), and leg pain scores were 0.5 and 2.1 ($P = 0.001$) for the solid fusion and pseudarthrosis groups, respectively.

Before surgery, spondylolisthesis measured 6.4 mm in the solid fusion group and 6.9 mm in the pseudarthrosis group, (range 2–18 mm). Preoperative sagittal motion averaged 3 mm for both groups (range 0–11 mm). Preoperative angular motion was 6.6° for the arthrodesis group and 10.1° for the pseudarthrosis group (range 0–17°).

After surgery, the spondylolisthesis averaged 7.3 mm and sagittal and angular motion decreased to 2.6 mm and 8.4°, respectively, in the pseudarthrosis group. The amount of spondylolisthesis remained the same for the solid fusion group, whereas sagittal and angular motion decreased to 1.0 mm and 0.5°, respectively. The significant improvement in postoperative dynamic instability in the solid fusion group is a product of the arthrodesis. The slip severity and sagittal and angular motion were then analyzed to determine what effect, if any, they have on the likelihood of a solid fusion occurring.

Preoperative angulation averaged 6.6° in those patients who eventually had a solid fusion, compared with 10.1° in those in whom a pseudarthrosis developed ($P = 0.02$, Student's t test of independent samples). This difference between the groups was statistically significant. Preoperative spondylolisthesis ($P = 0.66$) and sagittal motion ($P = 0.89$) were not predictive of fusion outcome.

Evaluation on the self-administered spinal stenosis questionnaire revealed that the solid fusion group scored statistically significantly better in the symptom severity and physical function categories. There was no statistical difference between the solid fusion and pseudarthrosis group on the patient satisfaction scale.

There was no statistically significant difference discovered between the two groups with respect to the major influencing variables of age, sex, levels fused, smoking, diabetes, or peripheral vascular disease.

There were no new peripheral (lower motor neuron) neurologic deficits after surgery in either group. No postoperative infections developed.

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

Seven patients underwent a second lumbar spine surgery after the original index procedure. Five of these patients were of the pseudarthrosis group and 2 of the arthrodesis group. Three patients with pseudarthrosis, all with poor results, went on to have a second attempt at arthrodesis, this time with instrumentation. The remaining 2 patients from the pseudarthrosis group both underwent decompressive lumbar laminectomy at a spinal location different from the original surgery. Of the 2 patients who had a solid fusion and underwent a second lumbar spine procedure, both had decompressive lumbar laminectomy at a spinal location different from the original surgery.

Discussion

Nonoperative methods are effective in the treatment of most patients with symptomatic degenerative spondylolisthesis and spinal stenosis. Initial treatment consists of short-term activity restriction and a nonsteroidal analgesic, if tolerated. Physical therapy along with massage, heat, ultrasound, and limited pelvic traction may be used as well. Ultimately, patients are recommended to establish a regular exercise program consisting of aerobic, active flexion, and abdominal and back strengthening exercises. Surgery was advised to patients in this study who failed to respond to a reasonable trial of nonoperative treatment for a minimum of 3 months.

Surgical management of degenerative spondylolisthesis with spinal stenosis has evolved from decompressive lumbar laminectomy alone to decompression combined with a fusion procedure. The clinical benefits of performing an arthrodesis following decompression have been substantiated by several studies.[5,11–13,23,24] In addition, there is justification from the literature to promote the use of spinal instrumentation as a means to increase the fusion rate.[12–16,18,25] At the same time, good to excellent clinical outcomes have been demonstrated independent of the radiographic fusion status.[18] The long-term status and the implications of a pseudarthrosis and its association vis-à-vis clinical outcome has not yet been elucidated.

Fischgrund *et al* reported on a study of 68 patients with degenerative spondylolisthesis and spinal stenosis.[18] This was a prospective, randomized study comparing decompressive lumbar laminectomy and arthrodesis with and without instrumentation. Thirty-five patients received pedicle screw instrumentation as part of their protocol and 33 did not. Whereas the results of the study did demonstrate an increased fusion rate with instrumentation (83% *vs.* 45%), there was no statistical difference in clinical outcome between treatment groups, after a minimum follow-up period of 24 months.

Bridwell *et al* performed a prospective randomized study of 49 patients with symptomatic degenerative spondylolisthesis.[13] There was an average follow-up of 3 years. Three treatment groups were established: Group 1, no arthrodesis; Group 2, decompression and arthrodesis; and Group 3, decompression and arthrodesis combined with pedicle screw instrumentation. An exception to the randomization process was made for those patients with >10° of angular motion or >3 mm of sagittal motion. This group of patients was automatically assigned to Group 3 and received instrumentation as part of their operative procedure. The results of this study demonstrated an improved fusion rate in the instrumented fusion group (87% *vs.* 30%) when compared with the noninstrumented fusion group. Functional assessment was determined by a single parameter, walking ability. Eighty-three percent of the instrumented fusion group felt that they were able to walk significantly better after the surgery compared with 31% reported for Groups 1 and 2.

Mardjetko *et al*, in 1994, published a meta-analysis of 25 publications and 889 patients with degenerative spondylolisthesis, collected between the years 1970 and 1993.[14] Patients in these studies were categorized based on treatment methods, including decompression, fusion, non-pedicle screw instrumentation, pedicle screw instrumentation, and anterior fusions. Those patients treated with decompression alone reported a 69% satisfactory clinical outcome. With the addition of an arthrodesis, the satisfaction rate increased to 90% and the successful fusion rate was 86%. The combination of decompression, arthrodesis, and pedicle screw instrumentation revealed that 86% had satisfactory outcomes and the fusion rate was increased to 93%. However, there was no statistical significance when comparing the outcome results and fusion rates between the instrumented and noninstrumented fusion groups.

Zdeblick, *in 1993*, prospectively evaluated 124 patients undergoing lumbar or lumbosacral fusion for degenerative conditions of the spine.[16] All patients were randomized to one of three treatment groups. Noninstrumented fusions were compared with fusions with semirigid and rigid instrumentation systems. This study revealed improved clinical outcome (95% *vs.* 71% good to excellent results) and better fusion rates (95% *vs.* 65% successful fusions) in the rigidly instrumented patients compared with the noninstrumented fusion group. Only 26 of these patients were noted to have degenerative spondylolisthesis.

Yuan *et al* presented a historical cohort study of pedicle screw fixation in thoracic, lumbar, and sacral spinal fusions in 1994.[15] A total of 3,498 patients were included in this study, 2,684 of whom had degenerative spondylolisthesis. The majority of patients with degenerative spondylolisthesis (81%) were treated with pedicle screw instrumentation and autogenous arthrodesis. The pedicle screw group demonstrated a higher fusion rate (89% *vs.* 70%), improved spinal alignment, and a shorter time to fusion consolidation when compared with the noninstrumented fusion group. The pedicle screw treatment cohort also demonstrated improved clinical outcomes with better function, greater neurologic recovery, and less back and leg pain than the noninstrumented fusion control group. The authors con-

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

cluded that the clinical benefits of pedicle screw instrumentation in this condition outweigh any potential risk from implant breakage or other untoward perioperative event.

Booth *et al* reported on the clinical and radiographic outcome of 41 cases of degenerative spondylolisthesis treated with decompression and instrumented posterior fusion.[12] The authors demonstrated a satisfaction rate of 83% at final clinical evaluation. Eighty-six percent of patients reported a reduction in back or leg pain from their preoperative condition. Functional improvement after surgery was significant at the 2-year follow-up, but not at final review. There were no patients with a symptomatic pseudarthrosis. This was a retrospective review with a minimum 5-year follow-up (mean 6.5 years). No control group was available for comparison with the study cohort. Back and leg pain questions were grouped together rather than assessed apart. Eight patients had multiple-level fusions for adjacent segment subluxations. However, this study had been the longest follow-up of patients treated operatively for this disorder. It demonstrated that approximately 85% of patients treated in this manner, all of whom had a solid fusion, will maintain a satisfactory clinical outcome even after 5 years.

The current series is the longest prospective study of degenerative spondylolisthesis treated with decompression and arthrodesis. All patients were treated with single-level decompression and bilateral posterolateral autogenous fusion for degenerative spondylolisthesis with concurrent spinal stenosis. A successful arthrodesis was shown to generate improved long-term clinical results over pseudarthrosis. Clinical outcome was good to excellent in 86% of the patients with a solid fusion compared with 56% of patients with a pseudarthrosis ($P = 0.01$). Back and lower limb pain scores were statistically significantly improved as well. In short review, good to excellent results have been reported in patients despite a pseudarthrosis. These results, as shown in the current study, have not been maintained over time. Long-term clinical benefits of an arthrodesis over pseudarthrosis, with respect to back and lower leg symptomatology, are realized on later review.

A major difficulty encountered in this study was locating patients, some of whom had surgery as early as 1985. The average age at the time of the index surgical procedure was 72 years. Currently, the mean age of the study group is more than 80 years, with many of the participants having relocated to warmer climates. It was felt by the authors that it would be impossible to do a clinical and radiologic follow-up on this population due to geographic constraints. It has been assumed for this study that the patient's radiographic status at 3 years has been maintained over the course of the study. Therefore, if a patient demonstrated a pseudarthrosis at final radiologic follow-up (2–4 years), it would be unlikely that a solid arthrodesis would occur in the ensuing 5 to 10 years. If clinical and radiographic data were required on each patient to complete this study, we think that the attrition rate would be unacceptably high, therefore invalidating any results.

A successful arthrodesis correlates with better radiologic parameters as well as an improved clinical outcome. In this study, the solid fusion and pseudarthrosis groups had similar preoperative demographics. Preoperative radiographs were analyzed in an attempt to identify radiographic measures of spondylolisthesis severity, which may influence fusion outcome. The initial spondylolisthesis and sagittal motion were not predictive of radiographic fusion. The initial preoperative angular motion at the location of the spondylolisthesis was statistically higher in those patients who ultimately went on to pseudarthrosis. In 22 patients, in whom a solid fusion was achieved, the preoperative angulation averaged 6.6°, whereas angulation in the 25 patients in whom a pseudarthrosis developed averaged 11° before surgery ($P = 0.02$).

Conclusion

The results of the current study demonstrate that in patients undergoing single-level decompression and posterolateral arthrodesis for spinal stenosis for concurrent spondylolisthesis, a solid fusion provides lasting long-term clinical benefits. A successful fusion correlates with an improved functional outcome and less back and lower limb symptomatology, compared with prior shorter-term studies, which indicated no significant difference between the successful fusion and pseudarthrosis groups. An increased angular motion may be a preoperative marker for those patients at risk for the development of pseudarthrosis. The amount of preoperative spondylolisthesis and sagittal motion did not correlate with radiographic fusion status. Based on previous work, the addition of spinal instrumentation in this patient population increases the ability to obtain a solid fusion and may be recommended as an adjunct to bone grafting alone in patients at risk for pseudarthrosis.

Key Points

- Patients with spinal stenosis and degenerative spondylolisthesis benefitted from a solid arthrodesis.
- Patients who had a pseudarthrosis had an inferior long-term outcome.
- These results differ from previous shorter-term studies.

References

1. Newman PH. Spondylolisthesis: its cause and effect: Hunterian Lecture. *Ann R Coll Surg*. 1955;16:305–323.
2. Junghanns H. Spondylolisthesen ohne Spalt im Zwischergelenkstuclz ('Pseudospondylolisthen'). *Arch Orthop Unfallchirurgie*. 1931;29:118–127.
3. MacNab I. Spondylolisthesis with an intact neural arch: the so-called pseudo-spondylolisthesis. *J Bone Joint Surg Br*. 1950;32:325–333.
4. Wiltse LL, Newman PH, Macnab I. Classification of spondylolysis and spondylolisthesis. *Clin Orthop*. 1976;117:23–29.

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

5. Caputy A, Luessenhop A. A long-term evaluation of decompressive surgery for degenerative lumbar stenosis. *J Neurosurg*. 1992;77:669–676.
6. Cauchoix J, Benoist M, Chassaing V. Degenerative spondylolisthesis. *Clin Orthop*. 1976;115:123–129.
7. Dall B, Rowe D. Degenerative spondylolisthesis: its surgical management. *Spine*. 1985;10:669–672.
8. Herron LD, Trippi AC. The result of treatment by decompressive laminectomy without fusion. *Spine*. 1989;14:534–538.
9. Postacchini F, Cinotti G. Bone regrowth after surgical decompression for lumbar spinal stenosis. *J Bone Joint Surg Br*. 1992;74:862–869.
10. Rosenberg NJ. Degenerative spondylolisthesis. *Clin Orthop*. 1976;117:112–120.
11. Herkowitz HN, Kurz LT. Degenerative lumbar spondylolisthesis with spinal stenosis. *J Bone Joint Surg Am*. 1991;73:802.
12. Booth KC, Bridwell KH, Eisenberg BA, et al. Minimum five-year results of degenerative spondylolisthesis treated with decompression and instrumented posterior fusion. *Spine*. 1999;24:1721–1727.
13. Bridwell KH, Sedgewick TA, O'Brien MF, et al. The role of fusion and instrumentation in the treatment of degenerative spondylolisthesis with spinal stenosis. *J Spinal Disord*. 1993;6:461–472.
14. Mardjetko SM, Connoly PJ, Shott S. Degenerative lumbar spondylolisthesis: a meta-analysis of the literature 1970–1993. *Spine*. 1994;10(suppl):2256–2265.
15. Yuan HA, Garfin SR, Dickman CA, et al. A historical cohort study of pedicle screw fixation in thoracic, lumbar, and sacral spinal fusions. *Spine*. 1994;19(suppl):2279–2296.
16. Zdeblick TA. A prospective, randomized study of lumbar fusion. *Spine*. 1993;18:983–991.
17. Herkowitz HN, Abraham DJ. Degenerative lumbar spondylolisthesis. *Semin Spine Surg*. 1999;11:28–33.
18. Fischgrund JS, Mackay M, Herkowitz HN, et al. Degenerative lumbar spondylolisthesis with spinal stenosis: a prospective, randomized study comparing decompressive laminectomy and arthrodesis with and without spinal instrumentation. *Spine*. 1997;22:2807–2812.
19. Macnab I, Dall D. The blood supply of the lumbar spine and its application to the technique of intertransverse lumbar fusion. *J Bone Joint Surg Br*. 1971;53:628–638.
20. Wiltse LL, Bateman JG, Hutchinson RH, et al. The paraspinal sacrospinalis splitting approach to the lumbar spine. *J Bone Joint Surg Am*. 1988;50:919–926.
21. Wisneski RJ, Rothman RH. Posterior intertransverse process fusion: indications, pathomechanics and results. In: White AH, Rothman RH, Ray CD, eds. *Lumbar Spine Surgery: Techniques and Complications*. St. Louis: Mosby, 1987:2721–2745.
22. 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.
23. Feffer H, Wiesel S, Cuckler J, et al. Degenerative spondylolisthesis: to fuse or not to fuse. *Spine*. 1985;10:286–289.
24. Lombardi JS, Wiltse LL, Reynolds J, et al. Treatment of degenerative spondylolisthesis. *Spine*. 1985;10:821–827.
25. Rechtine GR, Sutterlin CE, Wood GW, et al. The efficacy of pedicle screw/plate fixation on lumbar/lumbosacral autogenous bone graft fusion in adult patients with degenerative spondylolisthesis. *J Spinal Disord*. 1996;9:382–391.

Point of View

Jeffrey N. Katz, MD, MS

Surgery for patients with functionally limiting pain arising from spinal stenosis associated with degenerative spondylolisthesis includes decompression of central and foraminal stenosis. In addition, there has been considerable debate over whether to include an arthrodesis and, if so, whether to augment the posterolateral arthrodesis with instrumentation. Two key studies have helped guide clinical decision-making in this area. In the controlled trial by Herkowitz and Kurz published over a decade ago, noninstrumented arthrodesis was superior to no arthrodesis in the management of spinal stenosis associated with degenerative spondylolisthesis.[1] This trial, and other evidence supporting arthrodesis, has prompted many surgeons to recommend arthrodesis in this clinical context. The randomized controlled trial of Fischgrund *et al*[2] addressed the question of whether the arthrodesis should be augmented with instrumentation, and found that patients receiving noninstrumented arthrodesis experienced similar levels of symptom relief and functional improvement as patients receiving instrumented arthrodesis. In this trial, patients who received instrumented arthrodesis had higher rates of solid fusion than those who received noninstrumented arthrodesis, but the technical success of the fusion (solid *vs*. pseudarthrosis) was not associated with the extent of pain relief or functional improvement. Given the greater costs and complications associated with instrumented fusion, formal cost-effectiveness analyses[3] have suggested that noninstrumented arthrodesis has acceptable cost-effectiveness ($56,000 per quality-adjusted life year) when compared with decompression without arthrodesis, while instrumented arthrodesis had unacceptable cost-effectiveness (over 3 million dollars per quality-adjusted life year). If, however, instrumented arthrodesis resulted in substantially better symptom relief and functional improvement than noninstrumented arthrodesis, its cost-effectiveness would improve dramatically.[3]

Both of these pivotal studies reported results after 2 years of follow-up. In this issue of *Spine*, Kornblum *et al* performed a longer-term follow-up, restricted to the patients in these two trials who received noninstrumented arthrodesis. They did not follow the patients in the Herkowitz trial who received decompression without fusion or the patients in the Fischgrund trial who received instrumented fusion. The authors examined radiographs

From the Division of Rheumatology, Immunology and Allergy, Brigham and Women's Hospital, Boston, MA.
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.
Federal and 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 to Jeffrey N. Katz, MD, Division of Rheumatology, Immunology and Allergy, Brigham and Women's Hospital, 75 Francis Street, Boston, MA 02115; E-mail: jnkatz@partners.org

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

obtained after about 3 years of follow-up to ascertain whether the fusion was solid. In phone interviews conducted 5 to 14 years after surgery, they determined the patients' levels of pain and functional status. The findings are strikingly different from those of the 2-year analyses. Whereas the technical success of the arthrodesis was not associated with pain relief and functional improvement after 2 years of follow-up,[2] patients who achieved a solid fusion had considerably lower levels of back and leg pain and better functional status than patients with a pseudarthrosis after 5 to 14 years of follow-up.

One important clinical implication of this study is that instrumentation may offer long-term benefits not seen in the short-term trials. This critical inference is speculative and could only be proven with a randomized trial that had long-term follow-up. Given the expense of mounting such trials, it is disappointing that the authors chose to follow just one treatment arm from their trials and not both. They had an opportunity to compare the long-term outcomes of decompression with and without arthrodesis, and of arthrodesis with and without instrumentation. Such comparisons would address directly the issue that this study raises implicitly: If higher rates of solid fusion are associated with better pain relief, is instrumented arthrodesis a better choice than noninstrumented arthrodesis, or than no arthrodesis at all? In the absence of such direct comparisons, we must be cautious. While the higher fusion rates afforded by instrumented arthrodesis might lead to less back and leg pain, it is also possible that the instrumentation (and even the bone graft harvesting) could cause bothersome symptoms that would vitiate the benefits of a solid fusion over time.

Thus, this paper informs but does not resolve the debate over whether to add instrumentation to an arthrodesis for spinal stenosis and associated degenerative spondylolisthesis. The answer will await controlled trials with long-term follow-up. The paper does beg the question of the mechanism for superior long-term pain relief associated with solid fusion. Further research in this direction may yield additional insights into this important clinical problem.

References

1. Herkowitz HN, Kurz LT. Degenerative lumbar spondylolisthesis with spinal stenosis: a prospective study comparing decompression with decompression and intertransverse process arthrodesis. *J Bone Joint Surg Am*. 1991;73:802–808.
2. Fischgrund JS, et al. 1997 Volvo Award winner in clinical studies. Degenerative lumbar spondylolisthesis with spinal stenosis: a prospective, randomized study comparing decompressive laminectomy and arthrodesis with and without spinal instrumentation. *Spine*. 1997;22:2807–2812.
3. Kuntz KM, et al. Cost-effectiveness of fusion with and without instrumentation for patients with degenerative spondylolisthesis and spinal stenosis. *Spine*. 2000;25:1132–1139.

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

DONOR SITE PAIN FROM THE ILIUM

A COMPLICATION OF LUMBAR SPINE FUSION

B. N. SUMMERS, S. M. EISENSTEIN

From the Robert Jones & Agnes Hunt Hospital, Oswestry

Chronic pain at the donor site was reported by 25% of 290 patients who had undergone anterior lumbar spine fusion for low back pain. Donor site pain has characteristic clinical features, may be severely disabling and is stubbornly resistant to treatment.

The highest prevalence was in patients who had a tricortical full thickness graft taken through a separate incision overlying the iliac crest. Patients with a clinically unsatisfactory result from the spine fusion also had a significantly higher prevalence of donor site pain.

The use of autogenous bone graft in orthopaedic practice is common, and the ilium provides a large and accessible source. Postoperatively, patients often have more pain from the donor site than from the primary operation. This pain usually resolves over a period of several weeks, but it may persist.

Other complications of iliac bone grafting have been reported, such as fracture of the wing of the ilium, herniation of abdominal contents, and meralgia paraesthetica (Reid 1968; Weikel and Habal 1977; Guha and Poole 1983), but little detailed information is available about donor site pain. Laurie et al (1984) reported such chronic pain in 10% of patients after iliac crest grafting for maxillofacial procedures and considered that preservation of the crest itself was important, but gave no proof. Cockin (1971), in a review of 118 orthopaedic patients, found only 6% with donor site pain, hypersensitivity or buttock anaesthesia but gave no indication of the nature of the graft taken, the surgical approach, or the characteristics of the pain.

Large blocks of corticocancellous bone are required for anterior interbody lumbar fusion, and are usually taken from the anterior aspect of the ilium. Anterior fusion was the most commonly performed operation in this department for severe low back pain secondary to internal disc disruption, facet arthrosis, spondylolisthesis, or failure of other spine operations and was often combined with a posterior fusion under the same anaesthetic.

During follow-up, a significant proportion of these patients complained of disabling donor site pain. We therefore investigated this complication to determine its prevalence, nature, and predisposing features.

MATERIALS AND METHODS

We sent a postal questionnaire about donor site pain to 428 patients aged from 17 to 62 years who had had an anterior spinal fusion for low back pain.

Of these, 290 (68%) replied and 81 of them, about half with and half without donor site pain, were assessed clinically by one of the authors (BNS). Fifty-eight of these patients agreed to have radiographs of the donor site.

Definition. Donor site pain was defined as pain, with or without paraesthesia, which was separate from, and independent of any residual low back pain.

Grading. From the replies to the questionnaire, patients were placed in one of three groups, according to their assessment of the severity of the donor site pain:

1) pain which constituted a significant and unacceptable disability,
2) pain which the patient considered to be an acceptable symptom, and
3) no pain.

Surgical technique. Bone graft was taken from the anterior two-thirds of the ilium, approached either through the abdominal incision used to expose the anterior lumbar spine, or through a separate incision overlying and parallel to the iliac crest. A full thickness graft was taken, either as a tricortical block, incorporating the crest

B. N. Summers, FRCS, Senior Orthopaedic Registrar
The Middlesex Hospital, Mortimer Street, London W1P 7PN, England.

S. M. Eisenstein, PhD, FRCS, Director
The Department for Spinal Disorders, The Robert Jones & Agnes Hunt Orthopaedic Hospital, Oswestry, Salop SY10 7AG, England.

Correspondence should be sent to Mr B. N. Summers.

© 1989 British Editorial Society of Bone and Joint Surgery
0301-620X/89/4145 $2.00
J Bone Joint Surg [Br] 1989;71-B:677-80.

(Fig. 1), or as a bicortical block, leaving the crest intact (Fig. 2). In 11 of the 290 patients, acrylic cement was used to fill the donor site defect at the initial operation.

Statistical analysis. Statistical analysis of the results was performed by the Department of Statistical Science, University College, London employing the chi-squared test.

RESULTS

Incidence. The questionnaire revealed 'significant' donor site pain in 25% of patients, 'acceptable' pain in 24%, and no pain in 51%.

Of 290 patients responding to the questionnaire, the notes and radiographs of 235 patients were available for study. This showed a higher incidence of donor site pain in patients who had had a tricortical graft taken through a separate incision ($p < 0.05$), and in those patients who had an unsuccessful clinical result as regards back pain ($p < 0.001$).

Acrylic cement. In 11 of the 290 patients, acrylic cement had been used to fill the donor site. Of these 10 complained of donor site pain, and six considered it a significant disability. Five of the 290 donor sites had become infected and four of these were in patients with local acrylic cement.

Clinical assessment. The results of clinical assessment of 81 patients are given in Table I. Patients with donor site pain were older and more likely to be female, compared with asymptomatic patients, but this difference was not statistically significant.

Surgical approach. There was significant donor site pain

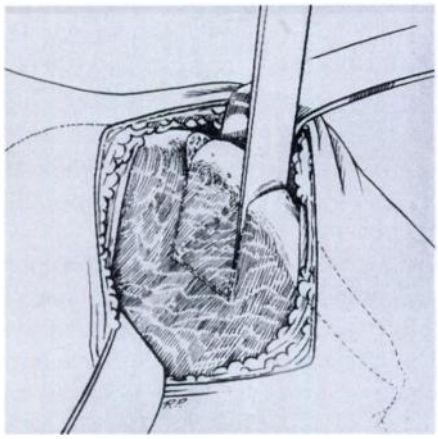


Fig. 1

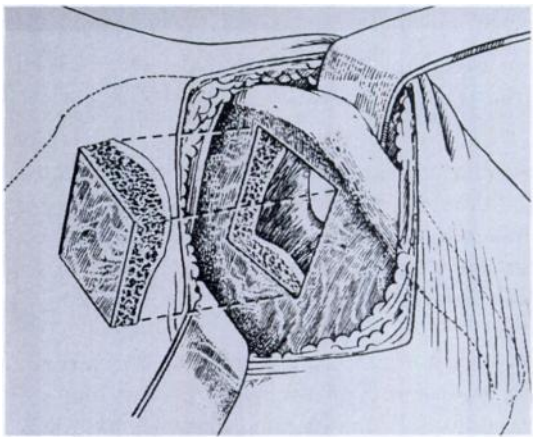


Fig. 2

Figure 1 - Operative removal of a tricortical graft including the iliac crest. Figure 2 - A bicortical graft preserving the crest.

Table I. Clinical details of the 81 patients who were examined

Donor site pain	Number of patients (*per cent* of total)	Donor sites number (*per cent*)	Males/females number (ratio)	Age at operation (years)	Follow-up (years)
Significant	25 *31*	26 *31*	7/18 1:2.6	43.2 21 to 62	5.4 1.8 to 9.8
Acceptable	14 *17*	14 *17*	8/6 1.33:1	38.1 17 to 51	4.9 0.8 to 10.3
None	42 *52*	42 *51*	19/23 1:1.21	34.2 17 to 52	4.7 0.7 to 10.2
All patients	81	82	34/47 1:1.38	37.6 17 to 62	5 0.7 to 10.3

Table II. Surgical approach and type of graft related to donor site pain in 81 examined patients with 82 donor sites (number and *per cent*). Surgical approach uncertain at nine donor sites

Donor site pain	Abdominal incision tricortical graft	Iliac incision	
		Tricortical	Bicortical
Significant	9 *30*	7 *88*	8 *33*
Acceptable	4 *13*	1 *12*	6 *17*
None	17 *57*	0	21 *60*
All patients	30	8	35

Table III. Clinical success of the spine operation related to donor site pain in 81 examined patients with 82 donor sites (number and *per cent*)

Donor site pain	Outcome of spine operation	
	Satisfied	Dissatisfied
Significant	5 *17*	21 *40*
Acceptable	4 *14*	10 *19*
None	20 *69*	22 *42*
All patients	29	53

Fig. 3

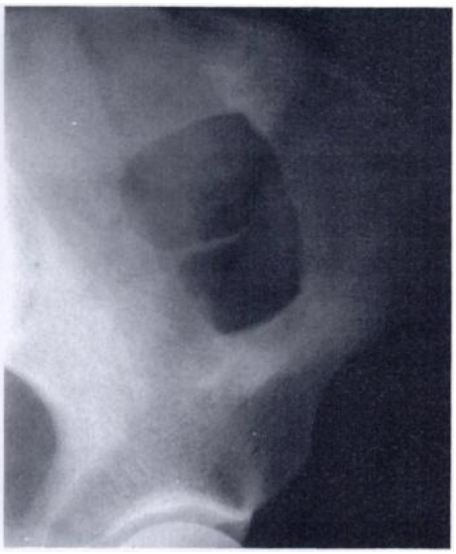


Fig. 4

Figure 3 - Radiograph of tricortical donor site four years after operation.
Figure 4 - Radiograph of a bicortical donor site six years after operation.

in 30% of the patients who had had a tricortical graft taken through an abdominal incision, 23% of those who had had a bicortical graft taken through a separate incision, and 88% of those who had had a tricortical graft taken through a separate iliac incision ($p < 0.05$ compared with the other two groups). Details are shown in Table II.

Success of the spinal fusion. Patients who considered the operation to have relieved their back pain suffered less donor site pain than patients who were dissatisfied with the clinical outcome ($p < 0.05$). Details are given in Table III.

Treatment. Treatment for donor site pain had included oral analgesia, local and epidural anaesthetic injections, local heat and ultrasound, cryotherapy, operative trimming of the iliac crest, immobilisation in a hip spica, and removal of acrylic cement. None of these treatments had relieved pain to any significant degree in any patient. Spontaneous resolution had occurred in only two patients at four and six years postoperatively.

Characteristics. Twenty-one of the 39 symptomatic 'examined' patients (54%) described the pain as burning, aching or of a 'toothache' nature, and nine (23%) as sharp or shooting, while seven (18%) described both these qualities of pain. Three patients (8%) were unable to describe their pain.

Thirty-four of the 39 symptomatic patients (87%) were unable to lie on the affected side because of pain. Twenty-seven patients (69%) found that pain was aggravated by walking, 14 (36%) by sitting and four (10%) by standing. After operation, pain had been immediate in 27 patients (69%), had developed between one and three months later in six patients (15%), and between four and 12 months later in the remaining six patients (15%). Only two of the symptomatic 'examined' patients had a positive Trendelenberg sign on the affected side. One of these patients also had tuberculosis of the hip, which was thought to account for this finding.

The site of maximal pain was over the iliac crest in 16 (76%) of 21 examined patients who had had tricortical grafts; and over the iliac wing in 14 (88%) of 18 examined patients who had had bicortical grafts. In the four remaining patients it was difficult to determine the maximal site of pain.

Radiography. Radiographs in 58 patients showed that none of the donor defects had filled in with new bone to any significant degree. Sharp edges had tended to round off, and some bony spikes were seen to project into the defect (Figs 3 and 4).

DISCUSSION

We found a high incidence of donor site pain in our patients. The pattern was characteristic, being maximal over the donor site, and usually aching, burning or 'like toothache'. It was typically aggravated by local pressure, and by walking, and many of the examined patients were unable to lie comfortably on the affected side. The pain was easily distinguished as being different in nature and site from any residual low back pain.

Bicortical grafts, preserving the crest itself, or tricortical grafts whether taken through the abdominal incision or through a separate one, all gave rise to a significant incidence of pain. The prevalence, however, was considerably higher after tricortical grafts taken through a separate incision and we presume that the close proximity of the scar to the donor defect, lying just subcutaneously, accounted for this finding. We therefore recommend that the crest itself is preserved unless it is approached through the existing abdominal incision.

The incidence of donor site pain was substantially higher in patients who considered that the fusion had not

relieved their back pain, indicating a possible psychological element in the perception of donor site pain.

The use of acrylic cement to fill the donor site was almost invariably associated with donor site pain and in addition there was a very high incidence of local infection. Clearly, this attempt at 'prophylaxis' was unsuccessful.

The precise cause of donor site pain remains obscure. We can postulate that it is either muscular or periosteal, secondary to the stripping of the abductors from the ilium, or neurogenic secondary to sensory nerve injury. Some patients have no pain, while others have severe symptoms from the same approach and technique; this may indicate that in some a sensory nerve was damaged, but in others it fortuitously escaped. A proportion of patients did not develop donor site pain until some months after the operation. This suggests that the pain was delayed until the development of a neuroma, but the facts that pain was closely related to the exact position of the donor site, and was typically aggravated by walking, indicate a local muscular or periosteal origin. It may well be that the cause is multifactorial.

One of the most salient features of donor site pain was its stubborn resistance to treatment. None of the conventional treatments we used were successful, and spontaneous resolution was very rare. Surgeons performing anterior spinal fusion for low back pain, must take into account the risk of significant pain from donor sites and balance it against the benefits of using autogenous bone.

We are grateful for the assistance of Dr S. Galliven, PhD, of the Department of Statistical Science, University College, Gower Street, London, for the statistical assessment of our results.

We also acknowledge the help of Mrs H. Evans, Miss A. Davis, AAMS, and Miss R. Dutton, in the preparation of the typescript; Mr R. Pearson for artwork; and the Medical Photographic Departments at the Robert Jones & Agnes Hunt Orthopaedic Hospital, Oswestry, and the Royal National Orthopaedic Hospital, Stanmore.

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

REFERENCES

Cockin J. Autologous bone grafting: complications at the donor site. *J Bone Joint Surg [Br]* 1971;53-B:153.

Guha SC, Poole MD. Stress fracture of the iliac bone with subfascial femoral neuropathy. Unusual complications at a bone graft donor site: case report. *Br J Plast Surg* 1983;36:305-6.

Laurie SWS, Kaban LB, Mulliken JB, Murray JE. Donor-site morbidity after harvesting rib and iliac bone. *Plast Reconstr Surg* 1984;73:933-8.

Reid EL. Hernia through an iliac bone-graft donor site a case report. *J Bone Joint Surg [Am]* 1968;50-A:757-60.

Weikel AM, Habal MB. *Meralgia paresthetica:* a complication of iliac bone procurement. *Plast Reconstr Surg* 1977;60:572-4.

Eur Spine J (2008) 17:1107–1112
DOI 10.1007/s00586-008-0695-9

ORIGINAL ARTICLE

Union versus nonunion after posterolateral lumbar fusion: a comparison of long-term surgical outcomes in patients with degenerative lumbar spondylolisthesis

Takahiro Tsutsumimoto · Mitsuhiko Shimogata · Yasuo Yoshimura · Hiromichi Misawa

Received: 3 December 2007 / Revised: 4 May 2008 / Accepted: 18 May 2008 / Published online: 7 June 2008
© Springer-Verlag 2008

Abstract It has been reported that in patients undergoing posterolateral lumbar fusion (PLF), the fusion status is not related to the short-term operative results. To determine whether the fusion status influences the long-term operative results of PLF, we retrospectively examined the surgical outcomes of uninstrumented PLF for a minimum of 8 years (average, 9.5 years), by comparing cases exhibiting union with those exhibiting nonunion. Uninstrumented PLF was performed for the treatment of lumbar canal stenosis (LCS) with degenerative spondylolisthesis. Since nine patients were lost to final follow-up, the study included 42 patients, and the follow-up rate was 82.4%. The mean age of the patients was 64.1 years (range 46–77 years). Eight patients exhibited fusion at the L3–4 level and 34 patients, at the L4–5 level. The fusion status was assessed using plain radiographs. The clinical outcomes were evaluated using the Japanese Orthopaedic Association (JOA) scores. Nonunion was noted in 26% (11/42) of the patients. There were no statistically significant differences between the groups exhibiting union and nonunion with respect to age, sex, preoperative JOA score, or preoperative lumbar instability. The union group achieved better operative results than the nonunion group at the 5-year and final follow-up ($P = 0.006$ and 0.008, respectively) although there was no significant difference in the percent recovery at 1 and 3-year follow-up ($P = 0.515$ and 0.506, respectively). A stepwise regression analysis revealed that the best combination of predictors for percent recovery at the time of final follow-up included the fusion status and the presence of comorbid disease. The results indicate that the fusion status following PLF is a critical factor influencing the long-term but not short-term operative results in the treatment of LCS with degenerative spondylolisthesis.

Key words Lumbar spinal fusion · Posterolateral lumbar fusion · Lumbar canal stenosis · Degenerative spondylolisthesis

T. Tsutsumimoto (✉) · M. Shimogata · Y. Yoshimura · H. Misawa
Spine Center, Yodakubo Hospital, 2857 Furumachi, Nagawa, Nagano 386-0603, Japan
e-mail: takatsutsumimoto@ybb.ne.jp

Introduction

Lumbar spine fusion is indicated as a primary procedure or as an adjunct to decompression for patients with degenerative spinal disorders for securing spinal stability and preventing postoperative instability [5, 17, 20]. Although several techniques of lumbar fusion exist, posterolateral lumbar fusion (PLF) is considered to be the gold standard for lumbar spinal fusion [1, 21]. PLF involves placing a bone graft harvested from the iliac crest between the transverse processes; this restricts spinal motion by bridging the posterolateral portion of the lumbar spine. In the surgical treatment of lumbar canal stenosis (LCS) associated with degenerative spondylolisthesis, a randomized trial and a study with alternating treatment assignments indicated better outcomes for laminectomy plus PLF than for laminectomy alone [7]. Fusion as an adjunct prevents the progression of spondylolisthesis after decompression and improves operative results possibly due to decreased postoperative back pain by the elimination of instability [7, 14]. These results have prompted many surgeons to recommend concomitant spinal fusion in the management of LCS associated with degenerative spondylolisthesis.

Nonunion following PLF may lead to changes in alignment, spinal instability and potential neurological injury;

however, asymptomatic nonunion has also been documented. In a 3-year prospective study comparing laminectomy alone to laminectomy combined with PLF in the treatment of LCS with degenerative spondylolisthesis, Herkowitz and Kurz [7] reported that the nonunion rate of PLF was 36% as observed on plain radiographs and that the clinical results were excellent or good for all patients who underwent PLF, including those demonstrating nonunion. In another report, Fischgrund et al. [6] published a 2-year prospective randomized study comparing the results of decompression and PLF alone with those of decompression and instrumented PLF in 67 patients with degenerative spondylolisthesis. This study demonstrated that in the uninstrumented PLF group, the clinical outcome was noted to be excellent or good in 83% of the patients who developed a pseudoarthrosis. Both studies indicated that the fusion status following uninstrumented PLF does not affect the short-term clinical outcome.

However, whether nonunion following uninstrumented PLF maintains good operative results on long-term follow-up remains an unresolved problem. In this study, we retrospectively studied the long-term operative results of uninstrumented PLF in patients having LCS with degenerative spondylolisthesis. The goals of the study were (1) to assess the outcome of uninstrumented PLF over time in terms of the fusion status and (2) to attempt to identify the demographic and clinical factors of the patients that were predictive of surgical outcomes.

Patients and methods

Patients

This study included patients who underwent decompressive surgery and a single-level PLF at the authors' institution for LCS with grade 1 degenerative spondylolisthesis. The patient exclusion criteria were as follows: (1) prior surgery to the lumbar spine, (2) isthmic spondylolisthesis, (3) segmental kyphosis at the listhetic segment and (4) completely collapsed disc height at the listhetic level. The charts of the patients who underwent the surgery between January 1995 and December 1997 were retrospectively reviewed. Spondylolisthesis was confirmed when the percentage of slip (%slip) was 5% or more on a lateral radiograph in a neutral standing position [15]. The total amount of angular motion between the adjacent vertebral endplate (dynamic angulation) and the total extent of the vertebral slip (dynamic translation) between the lateral radiographs taken in flexion and extension in the standing position were used to determine the dynamic motion at the fusion site [4]. In order to eliminate the X-ray magnification factor, the amount of translation was calculated as a percentage of the vertebral body width. Diagnosis was based on plain radiographic findings together with myelography, computed tomography (CT) and magnetic resonance imaging of the lumbar spine.

Operative technique

The decompression of the cauda equina and nerve roots was achieved by laminotomy (fenestration) or laminectomy with partial facetectomy (<50% on both sides). Discectomy was included when indicated. The transverse processes of the levels above and below the fusion segment were decorticated to expose the bone marrow. The autogenous iliac crest bone was placed so as to bridge the gap between the decorticated transverse processes. Concomitant spinal instrumentation was not used. Each patient was fitted with a lumbosacral brace and instructed to wear the brace when out of bed for 6 months following surgery.

Follow-up study

Radiographs (AP, oblique, lateral and flexion-extension) obtained at the final follow-up were examined to determine the fusion status of PLF. PLF was defined as union if radiographs demonstrated a bilateral continuity in the fusion mass between the cephalad and caudad transverse processes with less than 2° of angular motion and no translation between the vertebrae at the level of PLF on lateral flexion-extension radiographs [6, 7, 13]. The absence of continuity in the fusion mass at any point between the transverse process on one or both sides, greater than 2° of angular motion or any translation was considered a failure of fusion with nonunion [6, 7, 13]. The Japanese Orthopaedic Association's (JOA) scores for the assessment of low back pain were reviewed to evaluate the conditions before surgery and the clinical results at 1, 3 and 5 years after surgery and at the final follow-up. All the patients were followed up at each time interval. The JOA score comprises nine points assigned for subjective symptoms, six points for clinical signs and 14 points for the restriction of activity of daily living (ADL); the total score is thus 29 points [9]. Among the subjective symptoms, low back pain (LBP) and leg pain were also evaluated using the JOA score. Both scores ranged from 0 (indicating continuous severe pain) to 3 (no pain). The percent recovery of the JOA score that indicated the degree of normalization after surgery was calculated using the formula specified by Hirabayashi et al. [8] which was as follows:

Percent recovery (%) = [(postoperative JOA score − preoperative JOA score)/(29 − preoperative JOA score)] × 100.

The outcome was graded '4' for an improvement in the recovery rate of 75% or more, '3' for 50–74% improvement, '2' for 25–49% improvement and '1' for 24%

improvement or less and for when revision surgery was required.

Statistical analysis

All radiological data and clinical charts of these patients were retrospectively reviewed by examiners other than the treating surgeons in a blinded manner. The results are expressed as the mean ± SD. To determine what factors might be associated with the operative result at the final follow-up, stepwise regression analysis was used to determine the best multiple regression models of the postoperative percent recovery and potential predictors assessed. All variables with F values below 4 were excluded from the regression analysis. The factors included age at the time of surgery (continuous), gender (categorical: 1 = male, 2 = female), the severity of preoperative symptoms (preoperative JOA score; continuous), preoperative %slip (continuous), preoperative dynamic translation (continuous), preoperative dynamic angulation (continuous), postoperative %slip (continuous), fusion status (categorical: 1 = union, 2 = nonunion) and presence of comorbid diseases (categorical: 1 = present, 2 = absent). For each radiological and clinical parameter, the statistical differences between union and nonunion or those obtained before and after surgery were compared by using Fisher's exact probability test, Mann–Whitney test, or Wilcoxon signed-ranks test. Statistical analyses were performed with the StatView program (version 5.0; Abacus Concept Inc., Berkeley, CA). $P < 0.05$ was considered statistically significant.

Results

Between January 1995 and December 1997, 51 patients who underwent the surgery met the selection criteria. One patient died due to causes unrelated to the surgical procedure, and eight patients could not be located; therefore, these nine patients were excluded from the study. Thus, we retrospectively reviewed 42 patients (25 female and 17 male, follow-up rate: 82.4%) for an average follow-up period of 9.5 years (range 8–10 years). No new neurological deficits were observed after surgery. At the time of surgery, the mean age was 64.1 years (range 46–77 years). Eight patients exhibited fusion at the L3–4 level and 34 patients, at the L4–5 level. Of the 42 patients studied, 8 (19%) had a comorbid disease influencing their walking ability. The comorbid diseases included advanced osteoarthrosis of the hip or knee joint necessitating arthroplasty ($n = 4$), Parkinson's disease ($n = 2$), cervical or thoracic myelopathy ($n = 4$) and rheumatoid arthritis ($n = 1$). Three patients suffered from two comorbid diseases.

The averaged JOA score was 13.2 (range 3–20 points) before surgery and 23.5 (range 11–29 points) at the final follow-up. At the final follow-up, the percent recovery was greater than 3 in 69.0% (29/42) of the patients (Fig. 1). Of the six patients with a grade 1 percent recovery at the final evaluation, four patients belonged to the nonunion group and two patients belonged to the union group. Among them, two patients underwent a revision surgery at least 1-year after the initial operation. One patient in the union group suffered from recurrent leg pain due to lumbar disc herniation at the adjacent level below the fused segment 18 months after the surgery and subsequently underwent facetectomy and spinal instrumentation. The other patient with persistent LBP and leg pain underwent PLIF and spinal instrumentation at the same level due to nonunion.

Nonunion developed in 26.2% (11/42) of the patients (Table 1). There were no significant differences between the union and nonunion groups with regard to age, gender, fusion level, preoperative %slip, preoperative dynamic translation, preoperative dynamic angulation, preoperative JOA scores and number of patients with comorbid diseases. Figures 2 and 3 show the comparative clinical results for the two groups. At 1 and 3-year follow-up, there was no significant difference in the overall percent recovery between the two groups (union vs. nonunion: 3.5 ± 0.8 vs. 3.4 ± 0.7, $P = 0.515$, and 3.4 ± 0.8 vs. 3.1 ± 1.2, $P = 0.508$, respectively) (Fig. 2). However, the percent recovery in the union group was significantly better than that of the nonunion group at 5-year and final follow-up (union vs. nonunion: 3.5 ± 0.7 vs. 2.5 ± 1.0, $P = 0.006$, and 3.3 ± 0.9 vs. 2.2 ± 1.2, $P = 0.008$, respectively) (Fig. 2). The averages of the preoperative LBP score and the leg symptoms score did not differ significantly between the union and nonunion groups ($P = 0.361$ and 0.535, respectively). One year after surgery, a significant improvement in these scores was noted in both the groups

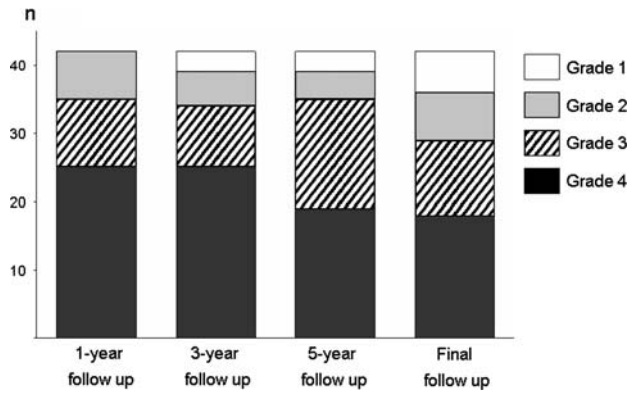


Fig. 1 Time-course changes in percent recovery ($n = 42$). The grades include grade 4 (percent recovery 75–100%), grade 3 (percent recovery 50–74%), grade 2 (percent recovery 25–49%) and grade 1 (percent recovery <25%, revision surgery)

Table 1 Patients' demographic data

	Union ($n = 31$)	Nonunion ($n = 11$)	P value
Mean age at surgery (years)	63.4	66.2	0.183
range	46–77	49–75	
Sex (female/male)	19/12	6/5	0.733
PLF level			
L3–4	6	2	>0.999
L4–5	25	9	
Preoperative percent slip	12.2	14.6	0.271
range	5.0–20.8	6.1–24.3	
Dynamic translation (%)	6.0	5.9	0.853
range	0–17.1	2.4–11.6	
Dynamic angulation (°)	8.3	7.9	0.698
range	1–17	0–20	
Preoperative JOA score	13.5	12.5	0.615
range	3–20	5–17	
Comorbid disease [no. (%)]	5 (16)	3 (27)	0.412

PLF posterolateral lumbar fusion, *JOA* Japanese Orthopaedic Association

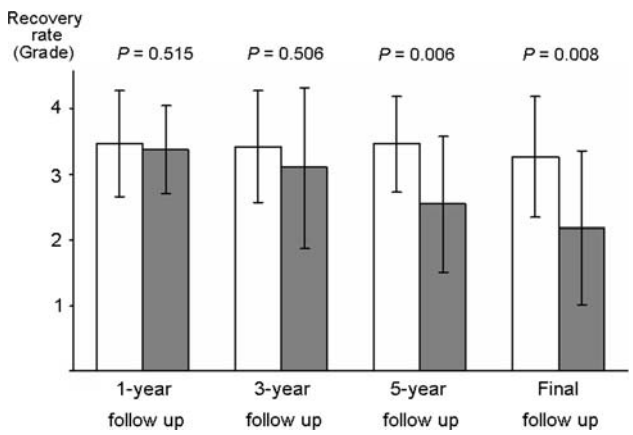


Fig. 2 Average percent recovery (grade) at different follow-up time points in the union and the nonunion groups. *Open bars*, union group ($n = 31$); *filled bars*, nonunion group ($n = 11$). The *error bars* represent standard deviation. The P values were calculated by using the Mann–Whitney test

(LBP: $P < 0.001$ in the union group and $P = 0.005$ in the nonunion group; leg symptoms: $P < 0.001$ in the union group and $P = 0.034$ in the nonunion group). At the final follow-up, the LBP score and the leg symptoms score in the union group were significantly better than the scores in the nonunion group ($P = 0.018$ and 0.035, respectively) (Fig. 4).

A stepwise regression analysis revealed that the best combination of the predictors for the percent recovery at the final follow-up included fusion status and the presence

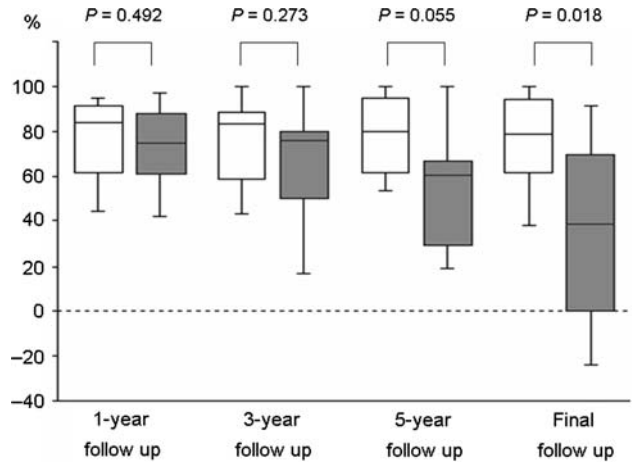


Fig. 3 Box plots showing the time-course changes in percent recovery (%) in the union and the nonunion groups. Patients who underwent revision surgery were excluded. *Open boxes*, union group ($n = 30$); *filled boxes*, nonunion group ($n = 10$). These plots illustrate the mean value as a *horizontal line* in a box. The vertical limits of the box define the 25th and 75th percentiles, and the *error bars* denote the 10th and 90th percentiles. The P values were calculated by using the Mann–Whitney test

of comorbid disease. The multiple regression equations are as follows: percent recovery (grade) = 4.334 − 1.076 (fusion status: union = 1, nonunion = 2), (adjusted $R^2 = 0.172$; $P = 0.004$), and percent recovery (%) = 48.927 − 31.309 (fusion status: union = 1, nonunion = 2) + 30.149 (comorbid disease: present = 1, absent = 2), (adjusted $R^2 = 0.311$; $P = 0.0004$).

Discussion

A fundamental problem that arises while investigating spinal fusion is the lack of definitive methods for confirming solid fusion. The fusion status can by accurately evaluated only through surgical exploration and direct inspection of the fusion mass; however, these methods are impractical for routine use. CT scanning has become the preferred diagnostic imaging modality for evaluating spinal fusion. Carreon et al. [2] have demonstrated that the positive predictive value for solid fusion on CT scans was 89%, while that of nonunion was only 74% when PLF on both sides were not fused on fine-cut CT scans with reconstructions. Thus, CT evaluation does not seem to be very reliable for the diagnosis of nonunion. Additionally, due to the harmful effects of radiation exposure, CT is not currently used as a routine method for fusion-status evaluation in our hospital. Although plain radiography is not the best method for assessing the fusion status [11], plain radiographs, accompanied by those in the flexion and extension bending views, are commonly used for this purpose because they are relatively inexpensive and easy to

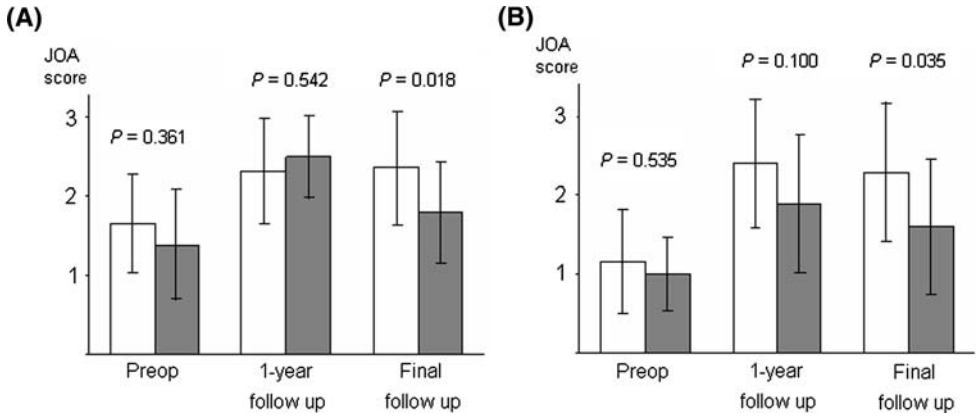


Fig. 4 Average scores of LBP (**a**) and leg symptoms (**b**) at different follow-up time points in the union and nonunion groups. *Open bars*, union group ($n = 30$); *filled bars*, nonunion group ($n = 10$). The *error bars* indicate standard deviation. The *P* values were calculated by using the Mann–Whitney test. Patients who underwent revision surgery were excluded. (*Preop* Preoperative values)

obtain [19]. Static plain radiography is used to detect the presence of a bone bridge between the transverse processes, and functional radiography is used to detect motion at the fused segment. In this study, the overall fusion rate achieved by PLF was 74%, as evaluated by the method described by Fischgrund et al. [6] and Kornblum et al. [13].

Even though the data suggest some beneficial effects of concomitant spinal fusion in the treatment of patients with degenerative spondylolisthesis [7], there is no agreement concerning the association between fusion status and surgical outcomes. In this study, we have retrospectively examined a minimum of 8-year surgical outcomes of decompression and PLF in the treatment of LCS with degenerative spondylolisthesis by comparing cases demonstrating union with those exhibiting nonunion. The results demonstrated that the union group achieved better clinical results than the nonunion group at the 5-year and final follow-up although no significant difference was observed at the 1 and 3-year follow-up. Additionally, the scores of LBP and leg symptoms in the union group were better than those observed in the nonunion group at the final follow-up, while these scores were not significantly different between the two groups at 1-year follow-up. In a 3-year prospective study comparing decompression alone with decompression and uninstrumented PLF in the treatment of patients with LCS and degenerative spondylolisthesis, Herkowitz and Kurz [7] reported that patients undergoing concomitant arthrodesis demonstrated improved clinical results, regardless of the fusion status, when compared with the 'decompression only' group. Thereafter, Kornblum et al. [13] described the long-term outcomes (mean, 7.7 years) of the patients treated with uninstrumented PLF in the studies by Herkowitz and Kurz [7] and Fischgrund et al. [6]. They demonstrated that patients exhibiting nonunion experienced significant deterioration in the surgical outcome as compared to a more stable long-term relief for patients with a solid fusion. These results together suggest that an arthrodesis attempt, regardless of the fusion status, appears to play a key role in the treatment of LCS with degenerative spondylolisthesis in the short term; however, the fusion status is a critical factor influencing the long-term operative results. The cause of deterioration in the long-term operative results in the nonunion group with time is a matter of debate. One possible explanation is that instability at the fusion segment in the nonunion group causes greater degenerative changes, such as laminar regrowth and hypertrophy of the facet joints, than that observed in the union group. These changes might lead to the recurrence of the spinal stenosis, resulting in the deterioration of surgical outcomes with time [3, 18].

Degeneration that develops at mobile segments above or below a fused spinal segment is known as adjacent segment disease (ASD) [16]. In this study, plain radiographs obtained at the time of final follow-up revealed degenerative changes in the regions adjacent to the fused segment in the case of eight patients (herniated nucleus pulposus in one patient, disc space narrowing in 4, and instability in 3) belonging to the union group. ASD did not seem to be related to the surgical outcomes, except in the case of one patient, who required revision surgery at the adjacent level, below the fused segment, due to lumbar disc herniation.

Regarding the predictors of surgical outcomes, a multiple regression analysis revealed that the coexistence of comorbid conditions was also a key predictor of the long-term operative results, which was consistent with previous studies [10, 12, 23]. This result is possibly caused by the influence of comorbid diseases on gait and ADL, which comprise 17 points in the JOA score. In contrast, a multivariate analysis was unable to identify a significant correlation between percent recovery and age, gender, preoperative %slip, preoperative JOA score, preoperative dynamic motion at the listhetic segment and postoperative %slip.

The present study has some limitations. First, it was conducted based on a set of retrospective data, and the outcomes were measured solely based on the JOA score

due to the unavailability of other scales such as the Visual Analogue Scale and the Oswestry Disability Index. Second, psychosocial factors which have been reported to influence the long-term surgical outcomes of spinal fusion [22] were not evaluated. Finally, the number of patients exhibiting nonunion was small ($n = 11$). Although the small sample size in the nonunion group clearly limited the results of the statistical analysis, we believe that this fact does not invalidate the main findings of our study.

Conclusions

In patients having LCS with degenerative spondylolisthesis who underwent uninstrumented PLF, the fusion rate evaluated by plain radiographs was 74%. The union group demonstrated better clinical results than the nonunion group in the long-term outcomes, while there were no significant differences in the short-term outcomes. Our results suggest that the fusion status in PLF and the coexistence of comorbid conditions are critical factors influencing the long-term operative results.

References

1. Boden SD (2002) Overview of the biology of lumbar spine fusion and principles for selecting a bone graft substitute. Spine 27:S26–S31. doi:10.1097/00007632-200208151-00007
2. Carreon LY, Djurasovic M, Glassman SD et al (2007) Diagnostic accuracy and reliability of fine-cut CT scans with reconstructions to determine the status of an instrumented posterolateral fusion with surgical exploration as reference standard. Spine 32:892–895. doi:10.1097/01.brs.0000259808.47104.dd
3. Chen Q, Baba H, Kamitani K et al (1994) Postoperative bone regrowth in lumbar spinal stenosis. A multivariate analysis of 48 patients. Spine 19:2144–2149
4. Dupuis PR, Yong-Hing K, Cassidy JD et al (1985) Radiologic diagnosis of degenerative lumbar spinal instability. Spine 10:262–276. doi:10.1097/00007632-198504000-00015
5. Esses SI, Huler RJ (1992) Indications for lumbar spine fusion in the adult. Clin Orthop:87–100
6. Fischgrund JS, Mackay M, Herkowitz HN et al (1997) 1997 Volvo Award winner in clinical studies. Degenerative lumbar spondylolisthesis with spinal stenosis: a prospective, randomized study comparing decompressive laminectomy and arthrodesis with and without spinal instrumentation. Spine 22:2807–2812. doi:10.1097/00007632-199712150-00003
7. Herkowitz HN, Kurz LT (1991) Degenerative lumbar spondylolisthesis with spinal stenosis. A prospective study comparing decompression with decompression and intertransverse process arthrodesis. J Bone Joint Surg Am 73:802–808
8. Hirabayashi K, Miyakawa J, Satomi K et al (1981) Operative results and postoperative progression of ossification among patients with ossification of cervical posterior longitudinal ligament. Spine 6:354–364. doi:10.1097/00007632-198107000-00005
9. Izumida S, Inoue S (1986) Assessment of treatment for low back pain. Japanese Orthopaedic Association. J Jpn Orthop Assoc 60:391–394
10. Jonsson B, Annertz M, Sjoberg C et al (1997) A prospective and consecutive study of surgically treated lumbar spinal stenosis. Part II: five-year follow-up by an independent observer. Spine 22:2938–2944. doi:10.1097/00007632-199712150-00017
11. Kant AP, Daum WJ, Dean SM et al (1995) Evaluation of lumbar spine fusion. Plain radiographs versus direct surgical exploration and observation. Spine 20:2313–2317. doi:10.1097/00007632-199511000-00009
12. Katz JN, Lipson SJ, Larson MG et al (1991) The outcome of decompressive laminectomy for degenerative lumbar stenosis. J Bone Joint Surg Am 73:809–816
13. Kornblum MB, Fischgrund JS, Herkowitz HN et al (2004) Degenerative lumbar spondylolisthesis with spinal stenosis: a prospective long-term study comparing fusion and pseudarthrosis. Spine 29:726–724. doi:10.1097/01.BRS.0000119398.22620.92
14. Mardjetko SM, Connolly PJ, Shott S (1994) Degenerative lumbar spondylolisthesis. A meta-analysis of literature 1970–1993. Spine 19:2256S–2265S. doi:doi:10.1097/00007632-199410151-00002
15. Nagaosa Y, Kikuchi S, Hasue M et al (1998) Pathoanatomic mechanisms of degenerative spondylolisthesis. A radiographic study. Spine 23:1447–1451. doi:10.1097/00007632-199807010-00004
16. Park P, Garton HJ, Gala VC et al (2004) Adjacent segment disease after lumbar or lumbosacral fusion: review of the literature. Spine 29:1938–1944. doi:10.1097/01.brs.0000137069.88904.03
17. Polly DW Jr, Santos ER, Mehbod AA (2005) Surgical treatment for the painful motion segment: matching technology with the indications: posterior lumbar fusion. Spine 30:S44–S51. doi:10.1097/01.brs.0000174529.07959.c0
18. Postacchini F, Cinotti G (1992) Bone regrowth after surgical decompression for lumbar spinal stenosis. J Bone Joint Surg Br 74:862–869
19. Resnick DK, Choudhri TF, Dailey AT et al (2005) Guidelines for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 4: radiographic assessment of fusion. J Neurosurg Spine 2:653–657
20. Sengupta DK, Herkowitz HN (2005) Degenerative spondylolisthesis: review of current trends and controversies. Spine 30:S71–S81. doi:10.1097/01.brs.0000155579.88537.8e
21. Tajima N, Chosa E, Watanabe S (2004) Posterolateral lumbar fusion. J Orthop Sci 9:327–333. doi:10.1007/s00776-004-0773-8
22. Trief PM, Ploutz-Snyder R, Fredrickson BE (2006) Emotional health predicts pain and function after fusion: a prospective multicenter study. Spine 31:823–830. doi:10.1097/01.brs.0000206362.03950.5b
23. Zheng F, Sandhu HS, Cammisa FP Jr et al (2001) Predictors of functional outcome in elderly patients undergoing posterior lumbar spine surgery. J Spinal Disord 14:518–521. doi:10.1097/00002517-200112000-00011

SPINE Volume 30, Number 24, pp 2709–2716
©2005, Lippincott Williams & Wilkins, Inc.

Comparison of OP-1 Putty (rhBMP-7) to Iliac Crest Autograft for Posterolateral Lumbar Arthrodesis

A Minimum 2-Year Follow-up Pilot Study

Alexander R. Vaccaro, MD,* D. Greg Anderson, MD,† Tushar Patel, MD,‡ Jeffrey Fischgrund, MD,§ Eeric Truumees, MD,‖ Harry N. Herkowitz, MD,¶ Frank Phillips, MD,# Alan Hilibrand, MD,** Todd J. Albert, MD,** Todd Wetzel, MD,†† and John A. McCulloch, MD‡‡

Study Design. A prospective, randomized, controlled, multicenter clinical study.

Objective. To compare the safety and clinical and radiographic outcomes of OP-1 (BMP-7) Putty to autogenous iliac crest bone graft in a population of patients undergoing laminectomy and posterolateral fusion for symptomatic lumbar stenosis associated with degenerative spondylolisthesis.

Summary of Background Data. Although the existing preclinical and clinical data suggest that OP-1 is able to achieve osteoinduction and clinical fusion in a variety of situations, the efficacy of this recombinant protein in a clinical spine fusion population has not been fully elucidated. This study directly compares the efficacy and safety of OP-1 putty to autograft bone for arthrodesis in patients with symptomatic stenosis in association with degenerative spondylolisthesis.

Methods. Thirty-six patients with degenerative lumbar spondylolisthesis and symptoms of neurogenic claudication underwent laminectomy, bilateral medial facetectomy, and posterolateral fusion using either iliac crest autograft or OP-1 Putty. Oswestry scores and SF-36 questionnaires were used to determine the clinical response to treatment. Independent, blinded neuroradiologists reviewed both static and dynamic radiographs to determine the fusion status. Successful fusion was declared when the presence of continuous bridging bone between the transverse processes was observed and less than 5° of angular motion and 2 mm of translational movement was measured using digital calipers.

Results. Efficacy data were tabulated for 27 patients at the 24-month time point and an additional 4 patients (without evaluable 24-month results) at the 36-month time point. One patient was not evaluable for radiology, so the data reflect clinical information for 31 patients and radiology for 30 patients. Clinical success, defined as a 20% improvement in the preoperative Oswestry score, was achieved by 17 of 20 (85%) OP-1 Putty patients and 7 of 11 (64%) autograft patients. A successful posterolateral fusion was achieved in 11 of 20 (55%) OP-1 Putty patients and 4 of 10 (40%) autograft patients. SF-36 scores showed similar clinical improvement in both groups. No systemic toxicity, ectopic bone formation, recurrent stenosis, or other adverse events specifically related to the use of the OP-1 Putty implant were observed.

Conclusion. This study represents the first clinical trial to demonstrate the safety and similarity of OP-1 Putty as a replacement for autogenous bone graft in the posterolateral fusion environment with a minimum of 2-year follow-up. OP-1 Putty was able to achieve osteoinduction leading to a radiographically solid fusion in the absence of autogenous iliac crest bone graft in 55% of the patients at 24 and 36 months. These results compare favorably to the historical fusion rates reported for uninstrumented arthrodesis in this challenging clinical scenario.

Key words: BMP, OP-1, fusion, posterolateral, lumbar spine, dynamic radiograph, degenerative spondylolisthesis. **Spine 2005;30:2709–2716**

From the *Department of Orthopaedics, Thomas Jefferson University, Rothman Institute, Philadelphia, PA; †Department of Orthopaedic Surgery, University of Virginia, School of Medicine, Charlottesville, VA; ‡Commonwealth Orthopaedics & Rehabilitation, Fairfax, VA; §William Beaumont Hospital, Royal Oak, MI; ‖Adjunct Faculty Bioengineering Center, Wayne State University, William Beaumont Hospital, Royal Oak, MI 48073; ¶Department of Orthopaedic Surgery, William Beaumont Hospital, Royal Oak, MI; #Department of Surgery, University of Chicago Spine Center, Chicago, IL; **Department of Orthopaedic Surgery, Thomas Jefferson University, Rothman Institute, Philadelphia, PA; ††Department of Orthopaedic Surgery, University of Chicago, Louis A. Weiss, MD, Memorial Hospital, Chicago, IL; and ‡‡Denver Orthopaedic Clinic, Denver, CO.

Acknowledgment date: August 3, 2004. First revision date: November 29, 2004. Acceptance date: January 14, 2005.

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 evaluating safety and efficacy of OP-1 (rhBMP-7) as a substitute for autologous bone graft in a posterolateral fusion.

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 reprint requests to Alexander R. Vaccaro, MD, Department of Orthopaedics, Thomas Jefferson University, Rothman Institute, 925 Chestnut St, 5th Floor, Philadelphia, PA; E-mail: alexvaccaro3@aol.com

Although posterolateral spinal arthrodesis is commonly used for the treatment of patients with symptomatic degenerative spondylolisthesis, fusion failure remains a common complication following surgery.[1,2] In addition to the lack of successful arthrodesis, donor site morbidity related to the bone graft harvest presents a problem that affects as many as 25% of patients following traditional spinal fusion using autogenous iliac crest bone graft.[3–5] A myriad of bone graft extenders and alternatives have been developed in an attempt to improve the rates of healing and avoid or diminish the complications of autograft harvest.[6–9] Unfortunately, none has yet proven effective enough to replace autograft bone as the gold standard for posterolateral spinal arthrodesis.

The discovery of osteogenic peptides by Marshall Urist in the mid 1960s heralded a new era in the science of bone formation and healing.[10,11] This family of re-

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

lated proteins has been subsequently named bone morphogenetic proteins (BMPs), and many members of this family have been isolated and characterized. Osteogenic BMPs have been shown to exert their action by recruiting and stimulating pluripotent mesenchymal cells along an osteoblastic lineage resulting in the formation of bone.[12] Because of the powerful osteogenic potential of these proteins, they have been studied with enthusiasm as a possible replacement or augmentation for autograft bone in spinal fusion. Several BMP preparations have been studied in preclinical and clinical trials for spinal applications.[13–18]

Osteogenic Protein 1 (OP-1), also called BMP-7, is a member of the TGF-β superfamily. Like other osteogenic BMPs, this protein induces the formation of bone when implanted in soft tissue ectopic locations. The human OP-1 gene has been cloned and introduced into a commercial cell line, facilitating the production of large quantities of recombinant human OP-1 (rhOP-1). OP-1 with various carrier preparations has been studied in a number of animal spine fusion models.[13–15,19–21] The available human data involving BMPs suggests that these molecules are associated with a low risk of protein related complications when used to promote bone healing or spinal fusion. Currently, OP-1 Putty has received a Humanitarian Device Exemption as of April 7, 2004, for use in the posterolateral spine by the United States Food and Drug Administration (FDA).

The purpose of this study was to assess the safety and efficacy of OP-1 Putty as a replacement for autograft bone when performing an uninstrumented posterolateral spinal arthrodesis in a population of patients with symptomatic lumbar spinal stenosis and degenerative spondylolisthesis. This study reports data for 36 patients randomized to either autograft or OP-1 Putty that were enrolled in a randomized, prospective, multicentered trial conducted under an Investigational Device Exemption Study as allowed by the U.S. FDA.

■ Materials and Methods

Study Design. The study was performed at five different institutions and involved 10 surgeons. At each site, Human Investigations Committee approval was obtained before patient enrollment. Thirty-six patients with single-level degenerative spondylolisthesis and stenosis at the L3–L4 or L4–L5 were enrolled. The patients were randomly selected to undergo a decompressive laminectomy and bilateral partial facetectomy combined with a posterolateral arthrodesis using either autologous iliac crest bone graft or OP-1 Putty. Patient randomization was done in a 2:1 fashion so that 24 patients received OP-1 Putty and 12 received autograft bone. Each patient was evaluated at the 6-week, and 3-, 6-, 9-, 12-, and 24-month time points following surgery, and yearly thereafter. Patient demographics are shown in Table 1. The surgical procedures were performed between June 1999 and January 2001.

Inclusion and Exclusion Criteria. All study patients had a Grade I or II degenerative spondylolisthesis of the L3–L4 or L4–L5 segments with coexistent spinal stenosis as confirmed by MRI or postmyelography CT. Clinically, the patients presented with symptoms of neurogenic claudication and had preoperative Oswestry scores of ≥30. All the patients were skeletally mature, and none had undergone previous lumbar surgery. All patients had failed at least 6 months of nonoperative treatment, including physical therapy, lumbar epidural injections, anti-inflammatory medications, and activity modifications for their spinal symptoms.

Patients with active spinal or systemic infection, a history of smoking, morbid obesity, or a known sensitivity to collagen were excluded from the study. Also excluded were pregnant women or those that planned to become pregnant. Patients with greater than 50% anterior translation of the cranial ver-

Table 1. Statistical Profile for All Patients, Including Age, Gender, Height, Weight, Fusion Level, and Preoperative Oswestry Score

	OP-1	Autograft	Significance
Age (yr) [mean (range)]	63 (43–80)	66 (51–79)	NS
Sex (female/male)	13/11	7/5	NS
Height (cm) [mean (range)]	169 (150–196)	177 (130–220)	NS
Weight (kg) [mean (range)]	74 (47–112)	66 (62–72)	NS
Level fused	L3–L4 = 1 (4%); L4–L5 = 23 (96%)	L3–L4 = 3 (25%); L4–L5 = 9 (75%)	NS
Preoperative Oswestry [mean (range)]	46 (16–68)*	47 (32–71)	NS

NS = not significant.
*One patient enrolled in the OP-1 arm was a protocol deviation because of the Oswestry score only being 16, and the minimum required score was 30. This was a calculation error, and the patient continued to be followed in the study.

Figure 1. **A**, A 24-month follow-up anteroposterior plain radiograph of an OP-1 patient showing solid bridging bone between the transverse processes. Lateral plain dynamic radiographs indicate essentially no translation in the sagittal plane on the (**B**) flexion or (**C**) extension views.

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

Table 2. Patients Evaluable for Clinical and Radiographic Endpoint Analyses

	24 Months		36 Months		Total	
	Clinical	Radiographic	Clinical	Radiographic	Clinical	Radiographic
Autograft	10	9	1	1	11	10
OP-1 Putty	17	17	3	3	20	20
Total	27	26	4	4	31	30

tebral body or greater than 20° of angular motion of the listhetic segment on flexion-extension films were excluded from study participation.

Randomization and Demographics. Patients were randomized in a 2:1 ratio to receive OP-1 Putty or iliac crest autograft for their spinal fusion. Randomization was performed after enrollment but before surgery using a computer algorithm (SAS using the PLAN procedure). Because of the study design, patients were not able to be blinded to the type of graft received.

In the OP-1 Putty group, 13 women (54%) and 11 men (46%) with an average age of 63 years (range, 43–80 years) were treated. Surgery was performed at L4–L5 in 23 patients (96%) and at L3–L4 in 1 patient (4%) (Table 1).

In the autograft group, 7 women (58%) and 5 men (42%) with an average age of 66 years (range, 51–79 years) were treated. Surgery was performed at L4–L5 in 9 patients (75%) and at L3–L4 in 3 patients (25%) (Table 1).

Fusion Materials. The OP-1 Putty implant consisted of 3.5 mg of rhOP-1 formulated with 1 g of Type 1 collagen derived from bovine bone and 200 mg of carboxymethylcellulose. This powdered mixture was reconstituted at the time of surgery by the addition of saline to achieve a final implant concentration of rhOP-1 protein of 0.875 mg/mL. One implant was used per side so that each patient received a total of 7 mg of rhOP-1 protein. No autogenous bone was used for the fusion in those patients randomized to receive the OP-1 Putty implants.

Patients who were randomized to the autograft group were treated with morselized corticocancellous bone harvested from the posterior iliac crest of the patient. No local bone graft was used for the fusion procedures. In both the OP-1 patients and the autograft patients, the implanted fusion material was placed between the decorticated transverse processes of the listhetic segment (*e.g.*, for a L4–L5 spondylolisthesis, the fusion material was used to bridge the space between the decorticated L4 and L5 transverse processes).

Surgical and Postoperative Protocol. All patients received a general anesthetic and prophylactic antibiotics. A posterior midline exposure was performed and carried out to the tips of the transverse processes of the listhetic segment. A bilateral laminectomy and bilateral medial facetectomies were performed to decompress the neural elements. The transverse processes of the levels above and below the slip were decorticated to expose the marrow elements of the bone. The fusion material (either one OP-1 Putty implant per side or half of the morselized autograft bone graft per side) was placed so as to bridge the space between the decorticated transverse processes. No irrigation was performed after placement of the fusion material.

Postoperative Management. Each patient was fitted with a rigid lumbosacral brace and instructed to wear the brace when out of bed for 3 months. Early ambulation was encouraged on the first day following surgery. Formal organized physical therapy emphasizing active exercises was begun 6 to 8 weeks following surgery. Each patient was scheduled for a follow-up visits with their surgeon at the 6-week, and 3-, 6-, 9-, 12-, and 24-month time points following surgery, and on a yearly basis after the 24-month visit. At each visit, a neurologic and radiographic assessment was performed. Oswestry and Medical Outcomes Study 36-Item Short Form Health Survey (SF-36) were administered at appropriate time points as described in the study protocol.

Radiographic Assessment. The radiographic assessment of the fusion status was the primary assessment for the effectiveness of the fusion material. This was determined by two independent neuroradiologists, blinded to the assigned treatment group of the patients. The radiologists studied anteroposterior, lateral, and flexion-extension lateral radiographs and measured the magnitude of the slip and angulation at the listhetic segment using digital calipers (Figure 1). The differences in slip and angulation measurement between flexion and extension lateral radiographs were used to determine motion at the fusion

Table 3. Success Rates: Clinical and Radiographic

	Success	% Success	95% CI (%)
Clinical			
OP-1	17/20	85	62.1–96.8
Autograft	7/11	64	30.8–89.1
$P = 0.21$			
Radiographic			
OP-1	11/20	55	31.5–76.9
Autograft	4/10	40	12.2–73.8
$P = 0.70$			

Note: Clinical success requires a 20%or greater improvement in Oswestry scores from pretreatment. Overall radiographic success, requiring an assessment of less than 5° of angular motion and less than 2 mm of translational movement on lateral flexion and extension radiographic views and bridging bone between the transverse processes on anteroposterior radiograph.

Table 4. Twenty-Four and Thirty-Six-Month Bridging Bone

Radiographic	Success	% Success	95% CI
OP-1	15/20	75	50.9–91.3
Autograft	8/10	80	44.4–97.5
$P = 1.00$			

Note: Bridging needs to be observed by 2 reviewers to be classified as a success.

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

Table 5. Adverse Event Table for the OP-1-Treated and Autograft-Treated Groups

Clinical	Adverse Events	% Adverse Events	95% CI
OP-1	23/24	96	78.9–99.9
Autograft	12/12	100	73.5–100.00
$P = 1.000$			

Table 7. Donor Site Pain at Various Follow-up Time Points for the Autograft Group

Visit Window	None (%)	Mild (%)	Moderate (%)	Severe (%)
6 wk	42	33	25	0
3 mo	27	55	18	0
6 mo	17	50	17	17
9 mo	22	33	44	0
12 mo	33	44	11	11
24 mo	33	22	44	0

site. To be considered a radiographically successful fusion, complete bridging bone had to be present between the transverse processes and less than 5° of angulation and ≤2 mm of translation had to be present at the site of the spondylolisthesis when comparing flexion to extension lateral radiographic views. If there was disagreement between the radiologists as to the fusion status of a patient, an evaluation by a third independent neuroradiologist was conducted and used as the definitive assessment.

Clinical Outcome. Clinical efficacy of the surgical procedure was measured using patient reported outcomes on the Oswestry questionnaire and SF-36 scale. The Oswestry score was used as the primary measure of clinical outcome with a clinically successful result arbitrarily defined as 20% improvement in the preoperative Oswestry score.

Safety and Adverse Events Reporting. The safety of the investigational product was evaluated by comparing the nature and frequency of adverse events in each of the two treatment groups. Adverse events included all minor and major medical events for which the patient sought medical attention regardless of the nature of the event or its severity.

Data Analysis and Statistics. Comparisons between treatment groups were analyzed using the two-sided Fisher's exact test with a P value of <0.05 considered statistically significant. Exact 95% confidence intervals were calculated for success rates.

■ Results

Patient Follow-up

As of June 2004, all patients still active in the study had passed the 36-month visit window and were being followed on a yearly basis. The study period was 24 months; after this visit, yearly long-term follow-up was scheduled. Before the 24-month visit, 3 patients in the OP-1 Putty group and 1 patient in the autograft group had discontinued the study: 3 either moved or were lost to follow-up and 1 patient had withdrawn voluntarily.

Efficacy Data Reporting (Radiographic and Clinical)

At the 24-month time point, 10 patients missed either the clinical or radiographic assessment (3 controls, 7 OP-1-treated patients). Of the 3 control patients with missing or incomplete data, 1 patient was lost to follow-up, 1 patient failed to have films taken at 24 months, and 1 patient failed to complete the Oswestry questionnaire. Of the 7 OP-1 patients who did not have complete data at 24 months, 1 patient had films taken but were poor in quality and found to be nondiagnostic by the independent radiologists, 1 patient failed to complete the Oswestry questionnaire, and 3 patients were lost to follow-up or withdrew from the study. The remaining 2 patients missed the 24-month assessment visit.

However, a complete set of 36-month clinical and radiographic data were available for 4 of the patients that did not have 24-month data: 3 patients in the OP-1 Putty group and 1 patient in the autograft group. These data are included with the minimum 24-month composite results. The 24-month clinical data from 1 autograft patient with inadequate radiographic studies was added to the reporting set for completeness of data information. A summary of this accounting is provided in Table 2.

Thus, the clinical data reported in this table consisted of 31 of 36 patients: 27 patients at 24 months and an additional 4 patients at 36 months. The radiographic data reported consisted of 30 of 36 patients, 26 patients at 24 months, and 4 additional patients at 36 months. Patients who failed to have clinical or radiographic data at either 24 or 36 months were excluded from the pri-

Table 6. No. of Reported Adverse Events by Treatment Group and Body System

	OP-1 Group (N = 24)			Autograft Group (N = 12)		
Body System	≥24 Months	Total No. of Events	% of All Events	≥24 Months	Total No. of Events	% of All Events
Cardiac	0	2	2.9	0	0	0
Gastrointestinal	0	2	2.9	1	3	6.7
Infections (superficial/deep)	0	5	7.1	0	1	2.2
Neural injury	0	2	2.9	1	3	6.7
Renal/urinary infection	3	4	5.7	0	2	4.4
Respiratory infection	0	1	1.4	0	0	0

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

Table 8. Presence of Straight Leg Tension Sign Causing Leg Pain*

Treatment	Preoperative	6 Weeks	3 Months	6 Months	9 Months	12 Months	24 Months†
OP-1 Putty	6/24 (25%)	3/24 (13%)	3/24 (13%)	3/24 (13%)	3/24 (13%)	1/22 (5%)	0/19 (0%)
Autograft	1/12 (8%)	0/12 (0%)	1/12 (8%)	1/12 (8%)	0/9 (0%)	1/11 (9%)	2/11 (18%)

*Pain at any angle.
†$P = 0.126$.

mary statistical analysis but are included in a worst-case analysis as described in a subsequent section.

Additional Data Reporting

Adverse Events and Complications

There were no medical complications or adverse events directly attributed to the OP-1 Putty product, with the exception of pseudarthrosis. The adverse events experienced by the patients enrolled in this study (Tables 3–5) were typical of the complications expected with lumbar decompression/fusion surgery given the age and diagnosis of the patient population as reported in Table 1. There was no significant difference in the rate of adverse events between the patients in either treatment group. Similarly, there was no significant difference noted between the operative times, hospital stays, or the presence of a postoperative straight leg tension sign between the two groups (Tables 6–12). No cases of systemic toxicity, ectopic bone formation, or recurrent spinal stenosis were observed in any of the patients in this study. No removals, revisions, or refusions with internal fixation had been performed in any of the patients in this study at the latest follow-up time point.

Table 9. Hospital Stay

Treatment	N	Mean	SD	Range
OP-1 Putty	23	3.9	1.7	2–10
Autograft	11	4.3	2.0	3–9

$P = 0.59$.

Table 10. Operative Time

Treatment	N	Mean	SD	Range
OP-1 Putty	24	138	43.0	50–220
Autograft	12	155	28.0	115–215

$P = 0.24$.

Radiographic Assessment

Using the strict criteria described above, radiographic fusion was observed in 11 of 20 (55%) OP-1 patients and 4 of 10 (40%) autograft patients at follow-up. This difference is not statistically significant (Table 13). Bridging bone between the transverse processes on the anteroposterior radiograph was observed in 75% of the OP-1 patients and 80% of autograft patients (Table 14).

Patient-Reported Clinical Outcomes: Oswestry

The Oswestry scores were used as the primary method of determining the clinical success of patients following surgery, with a clinical success arbitrarily defined as a 20% or greater improvement in their preoperative Oswestry score (Figure 2). At the 24- and 36-month visit, 17 of 20 (85%) OP-1 patients and 7 of 11 (64%) autograft patients achieved at least a 20% improvement in their Oswestry score and were graded as a clinical success. The difference in clinical success between autograft and OP-1-treated patients was not statistically significant. Figure 3 displays the degree of pain for the autograft group at various follow-up intervals.

Patient-Reported Outcomes: SF-36

The SF-36 survey measures the self-reported general well-being of a patient. SF-36 results are reported in Table 11–14 and Figure 4 (Figure 4). Both treatment groups showed an improvement in physical and mental well-being. Before surgery, the average Physical Component Summary scores for both treatment groups were below the 25th percentile of the normative data for age-matched citizens in the general U.S. population. At the 24-month visit, the mean Physical Component Summary

Table 11. SF-36 Scores Over Time: Physical and Mental Component Summary Scales

	Base	6 Weeks	3 Months	6 Months	9 Months	12 Months	24 Months
Autograft	30.1	31.7	36.6	35.9	35.8	38.3	29
OP-1	28.8	32.7	37.2	38.5	44	44.1	46.2

Table 12. SF-36 Scores Over Time: Norms for Physical Component Summary Scale in Selected Age Groups

Age (yr)	N	Mean (SD)	25th Percentile	50th Percentile	75th Percentile
55–64	269	45.90 (11.25)	38.66	49.86	54.32
65–74	442	43.33 (11.16)	35.04	46.18	52.50

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

Table 13. SF-36 Scores Over Time: Average Values of Mental Component Summary Scale

	Base	6 Weeks	3 Months	6 Months	9 Months	12 Months	24 Months
Autograft	44.4	50.4	51.2	52.6	50.5	52.8	50.4
OP-1	47.5	50.4	55.1	57.1	56.6	56.6	55.6

scores for the OP-1 treatment group were comparable to the mean age-matched normative values while SF-36 scores for the autograft treatment group were slightly lower than the mean age-matched normative values. Similarly, the mean Mental Component Summary scores for both treatment groups before surgery were comparable to the 25th percentile normative data for age-matched citizens in the U.S. population. At the 24-month visit, the mean Mental Component Summary scores for the OP-1 treatment group were between the 50th and 75th percentiles of the age-matched normative data, while the mean scores for the autograft group were slightly above the 50th percentile of the age-matched population controls.

A further calculation was performed on the assumption that patients lost to follow-up in both groups were clinical and radiographic failures and therefore chose not to return for further evaluation. With this assumption, 7 of 12 (58%) autograft patients would be rated a clinical success while 4 of 12 autograft patients (33%) would be rated a radiographic success. Similarly, in the OP-1 Putty group, 17 of 24 (71%) OP-1 patients would be rated a clinical successes while 11 of 24 (46%) would be rated a radiographic success. Using these assumptions, the differences between the clinical success rate and the radiographic fusion rate of the groups are still not statistically significant.

■ Discussion

In this study, the radiographic fusion rates and clinical outcomes have been carefully compared between two cohorts of patients undergoing posterolateral fusion using either OP-1 Putty or autograft iliac crest bone. At a minimum follow-up of 24 months, the rates of fusion and the clinical results were not statistically different between these two groups of patients. Also, the rates of adverse events were similar between the two groups, and the complications seen in both groups were typical for the age and diagnosis of the patients.

The posterolateral region of the spine is one of the more challenging fusion environments because of the large gap that must be spanned by bone, the relatively poor vascularity of this region, the tensile stresses present in this region of the spine, and the presence of motion when the fusion is performed without supplemental internal fixation.[22] Patients with the diagnosis of degenerative spondylolisthesis form a particularly challenging group, as these patients are generally older and have varying degrees of instability following decompression with partial facet removal.[1,2,22] In this population, Fischgrund *et al* observed a successful fusion rate of only 45% using autograft bone without supplemental internal fixation.[1] In the same study, internal fixation of the fusion site was associated with an increased fusion rate but did not increase the odds of a successful clinical result. For this reason, the use of internal fixation for the diagnosis of degenerative spondylolisthesis with stenosis remains controversial because the outcome data have not supported a dramatic benefit with the use of supplemental internal fixation, while the surgical costs and the risk of complications may be increased when internal fixation is used.[1,23–25]

Assessing the status of a posterolateral spinal fusion radiographically without open surgical exploration also presents a challenge. Although plain radiographs can be used to detect the presence of apparent bridging bone between the transverse processes, many studies have demonstrated the persistence of substantial motion suggesting failure to achieve a solid bony union.[1,2,26,27] Also, complicating the noninvasive assessment of posterolateral fusion is the fact that some motion occurs even in the setting of a solid posterolateral arthrodesis and the exact amount of motion that indicates a pseudarthrosis *versus* a fusion is unknown.[26] In this study, a very rigorous methodology was used to assess the status of the fusion. As mentioned, independent, blinded radiologists used digital calipers to assess the degree of motion and angulation on lateral flexion-extension radiographs. Both motion measurements and the presence of bridging bone on radiographs were necessary before classifying a patient as a fusion success. An important aspect of our assessment technique is the absence of internal fixation. Internal fixation can obscure the presence of bone; but more importantly, it can limit motion seen on flexion-extension films even in the setting of a pseudarthrosis. Although we believe that the current method for noninvasive assessment of the fusion is stringent compared with most published studies where lumbar fusion has been assessed, we acknowledge that no method, short of

Table 14. SF-36 Scores Over Time: Norms for Mental Component Summary Scale in Selected Age Groups

Age (yr)	N	Mean (SD)	25th Percentile	50th Percentile	75th Percentile
55–64	269	51.05 (9.69)	46.71	54.35	57.9
65–74	442	52.68 (9.29)	48.34	55.67	59.13

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

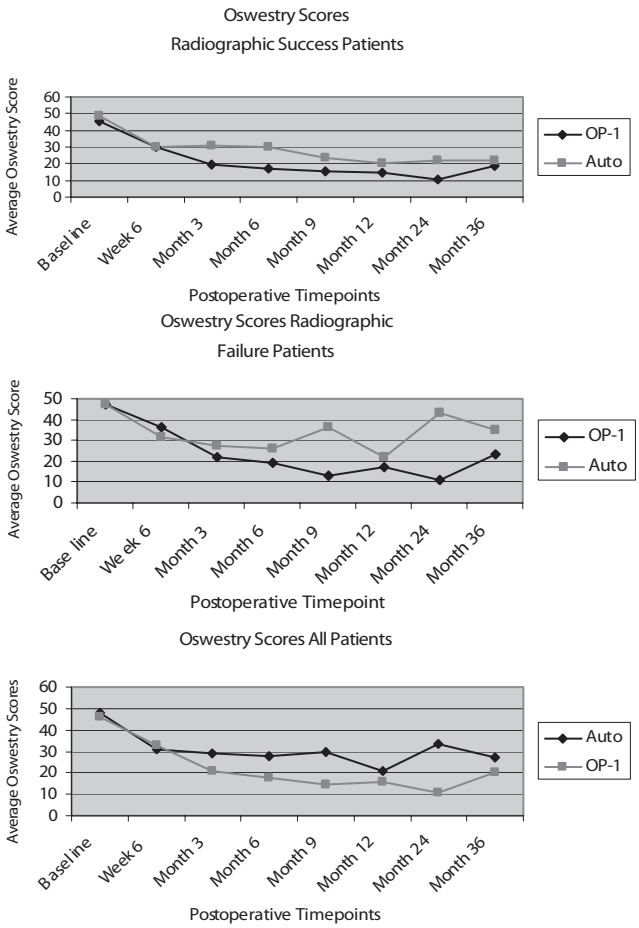


Figure 2. Oswestry scores. **A,** Radiographic success patients. **B,** Radiographic failure patients. **C,** All patients.

surgical exploration, is completely accurate for determining the status of a posterolateral fusion.[10] Eight patients (3 OP-1 Putty and 5 controls) in this study had apparent bridging bone between the transverse processes on anteroposterior radiographs but demonstrated ≥5° or more sagittal plane motion and were classified as fusion failures. Without surgical exploration, the true status of their fusions cannot be absolutely determined.

Animal and human studies done to date have suggested that BMPs used in clinically relevant doses are safe. Multiple studies have failed to demonstrate either systemic toxicity or tumor formation in response to these substances.[13–17,19–21] Even when rhOP-1 and a carrier were intentionally placed inside the dog dural sac, Paramore *et al* found no spinal cord inflammation, cytotoxicity, or tumorigenicity, although bone did form adjacent to the spinal cord.[19] Our study also supports the hypothesis that the OP-1 Putty implant is safe for use in humans undergoing decompression and fusion for degenerative spondylolisthesis as no evidence of systemic toxicity, ectopic bone formation, or implant migration into the laminectomy site was observed in this study.

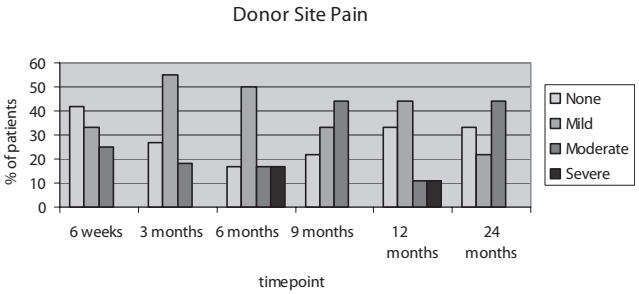


Figure 3. Donor site pain.

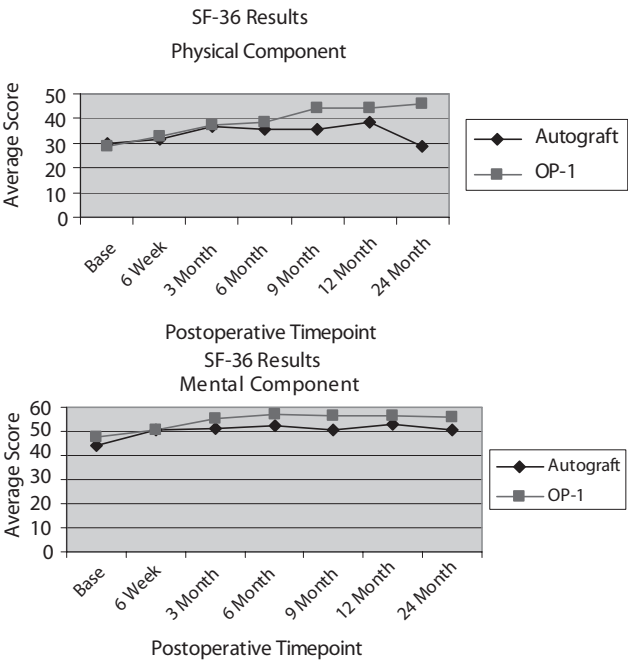


Figure 4. SF-36 results. **A,** Physical component. **B,** Mental component.

There are several limitations to this study that should be acknowledged. First, not all patients who volunteered to participate in the study were available at the 24- or 36-month follow up. Thus, this study should be classified as a pilot study with an incomplete ability to precisely determine the true statistical differences between the two treatment groups. With the availability of a larger sample size in the future, differences between the two groups could emerge despite the fact that no statistical differences were seen in the present study. Although the sample size was small, this limitation is compensated for to some degree by the inclusion of a single diagnosis and by ensuring good homogeneity between the treatment groups and surgical approaches. Another study limitation is the fact that not all enrolled patients were available to provide complete data at the 24-month time point. To compensate for this, 36-month data were added to the efficacy analyses for those patients not available for their 24-month evaluation. In addition, by including the worst-case analysis and assuming that all patients unavailable at the 24- or 36-month time point were clinical and radiographic failures, we have attempted to quantify the impact of this limitation on the study outcome. Another limitation is the absence of blinding on the part of either the patients or surgeons to the treatment group, which was not possible given the

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

design of this surgical study. However, the assessment of the fusion status was performed in a blinded fashion. Also, patient self-assessment outcomes were assessed using previously validated outcomes instruments.

Conclusion

Although the degenerative spondylolisthesis patient presents a challenging environment in which to study arthrodesis, this pilot study suggests that OP-1 Putty performs similarly to autograft bone when assessed radiographically and clinically. Overall, a successful radiographic fusion rate of 55% was achieved using OP-1 Putty, while the rate of bridging bone on the anteroposterior radiographs in this group was 75%. Clinically, 85% of the patients who underwent fusion using OP-1 Putty achieved a successful result from surgery (a 20% improvement in their preoperative Oswestry score). No signs of systemic toxicity, ectopic bone formation, or migration of the implant into the laminectomy site were observed in any patients in this study. The findings of this pilot study are in agreement with other published reports suggesting that BMPs have an acceptable safety profile for spinal fusion. Although this study appears promising, additional follow-up of these patients and further study with larger numbers of patients are required before firm conclusions can be reached as to the applicability of OP-1 Putty as an autograft substitute in posterolateral spinal arthrodesis.

Key Points

- OP-1 Putty was able to achieve osteoinduction, leading to a radiographically solid fusion in the absence of autogenous iliac crest bone graft in 55% of patients at 24 and 36 months, compared with a 40% fusion rate in autograph patients. This compares favorably with the historical fusion rates reported for uninstrumented arthrodesis.
- No systemic toxicity, ectopic bone formation, recurrent stenosis, or other adverse events specifically related to the use of the OP-1 Putty implant were observed.
- A total of 85% of the patients who underwent fusion using OP-1 Putty achieved a successful clinical result from surgery, defined as a minimum of 20% improvement in their preoperative Oswestry score.
- The rates of fusion and adverse events were similar between the OP-1 group and autograph iliac crest group (control), with complications seen in both groups that were typical for age and diagnosis.

References

1. Fischgrund JS, Mackay M, Herkowitz HN, et al. 1997 Volvo Award winner in clinical studies. Degenerative lumbar spondylolisthesis with spinal stenosis: a prospective, randomized study comparing decompressive laminectomy and arthrodesis with and without spinal instrumentation. *Spine* 1997;22:2807–12.
2. Herkowitz HN, Kurz LT. Degenerative lumbar spondylolisthesis with spinal stenosis: a prospective study comparing decompression with decompression and intertransverse process arthrodesis. *J Bone Joint Surg Am* 1991;73:802–8.
3. Beirne JC, Barry HJ, Brady FA, et al. Donor site morbidity of the anterior iliac crest following cancellous bone harvest. *Int J Oral Maxillofac Surg* 1996;25:268–71.
4. Cotler JM, Star AM. Complications of spinal fusion. In: Cotler JM, Cotler HB, eds. *Spine Fusions: Science and Technique*. New York: Springer-Verlag, 1990.
5. Younger EM, Chapman MW. Morbidity at bone graft donor sites. *J Orthop Trauma* 1989;3:192–5.
6. Boden SD. Biology of lumbar spine fusion and use of bone graft substitutes: present, future, and next generation. *Tissue Eng* 2000;6:383–99.
7. Hadjipavlou AG, Simmons JW, Tzermiadianos MN, et al. Plaster of Paris as bone substitute in spinal surgery. *Eur Spine J* 2001;2(suppl):189–96.
8. Li H, Zou X, Xue Q, et al. Anterior lumbar interbody fusion with carbon fiber cage loaded with bioceramics and platelet-rich plasma: an experimental study on pigs. *Eur Spine J* 2004;
9. Thalgott JS, Giuffre JM, Fritts K, et al. Instrumented posterolateral lumbar fusion using coralline hydroxyapatite with or without demineralized bone matrix, as an adjunct to autologous bone. *Spine J* 2001;1:131–7.
10. Urist MR. Bone: formation by autoinduction. *Science* 1965;150:893–9.
11. Van de Putte KA, Urist MR. Osteogenesis in the interior of intramuscular implants of decalcified bone matrix. *Clin Orthop* 1965;43:257–70.
12. Uludag H, Gao T, Porter TJ, et al. Delivery systems for BMPs: factors contributing to protein retention at an application site. *J Bone Joint Surg Am* 2001;83(suppl):128–35.
13. 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.
14. Grauer JN, Patel TC, Erulkar JS, et al. 2000 Young Investigator Research Award winner: evaluation of OP-1 as a graft substitute for intertransverse process lumbar fusion. *Spine* 2001;26:127–33.
15. 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.
16. Liao SS, Guan K, Cui FZ, et al. Lumbar spinal fusion with a mineralized collagen matrix and rhBMP-2 in a rabbit model. *Spine* 2003;28:1954–60.
17. Sidhu KS, Prochnow TD, Schmitt P, et al. Anterior cervical interbody fusion with rhBMP-2 and tantalum in a goat model. *Spine J* 2001;1:331–40.
18. Takahashi J, Saito N, Ebara S, et al. Anterior thoracic spinal fusion in dogs by injection of recombinant human bone morphogenetic protein-2 and a synthetic polymer. *J Spinal Disord Tech* 2003;16:137–43.
19. Paramore CG, Lauryssen C, Rauzzino MJ, et al. The safety of OP-1 for lumbar fusion with decompression: a canine study. *Neurosurgery* 1999;44:1151–5.
20. 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.
21. Magin MN, Delling G. Improved lumbar vertebral interbody fusion using rhOP-1: a comparison of autogenous bone graft, bovine hydroxylapatite (Bio-Oss), and BMP-7 (rhOP-1) in sheep. *Spine* 2001;26:469–78.
22. Boden SD. The biology of posterolateral lumbar spinal fusion. *Orthop Clin North Am* 1998;29:603–19.
23. Bridwell KH, Sedgewick TA, O'Brien MF, et al. The role of fusion and instrumentation in the treatment of degenerative spondylolisthesis with spinal stenosis. *J Spinal Disord* 1993;6:461–72.
24. Kuntz KM, Snider RK, Weinstein JN, et al. Cost-effectiveness of fusion with and without instrumentation for patients with degenerative spondylolisthesis and spinal stenosis. *Spine* 2000;25:1132–9.
25. Rechtine GR, Sutterlin CE, Wood GW, et al. The efficacy of pedicle screw/plate fixation on lumbar/lumbosacral autogenous bone graft fusion in adult patients with degenerative spondylolisthesis. *J Spinal Disord* 1996;9:382–91.
26. Barrick WT, Schofferman JA, Reynolds JB, et al. Anterior lumbar fusion improves discogenic pain at levels of prior posterolateral fusion. *Spine* 2000;25:853–7.
27. Heggeness MH, Esses SI. Classification of pseudarthroses of the lumbar spine. *Spine* 1991;16(suppl):449–54.

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

SPINE Volume 33, Number 26, pp 2850–2862
©2008, Lippincott Williams & Wilkins

The Safety and Efficacy of OP-1 (rhBMP-7) as a Replacement for Iliac Crest Autograft in Posterolateral Lumbar Arthrodesis

A Long-term (>4 Years) Pivotal Study

Alexander R. Vaccaro, MD, PhD,* James P. Lawrence, MD, MBA,* Tushar Patel, MD,† Lee D. Katz, MD,‡ D. Greg Anderson, MD,* Jeffrey S. Fischgrund, MD,§ Julie Krop, MD,¶ Michael G. Fehlings, MD,‖ and David Wong, MD**

Study Design. Randomized controlled trial comparing OP-1 (rhBMP-7) with iliac crest autograft in patients with symptomatic degenerative spondylolisthesis and spinal stenosis treated with decompression and uninstrumented posterolateral arthrodesis.

Objective. To determine the safety and the clinical and radiographic efficacy of OP-1 (rhBMP-7) Putty as compared with an iliac crest bone autograft control in uninstrumented, single-level posterolateral spinal arthrodesis.

Summary of Background Data. Preclinical and preliminary clinical data have demonstrated successful fusion and clinical outcomes with the use of OP-1 Putty in posterolateral spinal arthrodesis. No prior randomized controlled trial with adequate study power has been performed.

Methods. A total of 335 patients were randomized in 2:1 fashion to receive either OP-1 Putty or autograft in the setting of an uninstrumented posterolateral arthrodesis performed for degenerative spondylolisthesis and symptomatic spinal stenosis. Patients were observed serially with radiographs, clinical examinations, and appropriate clinical indicators, including ODI, Short-Form 36, and visual analog scale scores. Serum samples were examined at regular intervals to assess the presence of antibodies to OP-1. The primary end point, Overall Success, was analyzed at 24 months. The study was extended to include additional imaging data and long-term clinical follow-up at 36+ months. At the 36+ month time point, CT scans were obtained in addition to plain radiographs to evaluate the presence and location of new bone formation. Modified Overall Success, including improvements in ODI, absence of retreatment, neurologic success, absence of device-related serious adverse events, angulation and translation success, and new bone formation by CT scan (at 36+ months), was then calculated using the 24-month primary clinical endpoints, updated retreatment data, and CT imaging and radiographic end points.

Results. OP-1 Putty was demonstrated to be statistically equivalent to autograft with respect to the primary end point of modified overall success. The use of OP-1 Putty when compared to autograft was associated with statistically lower intraoperative blood loss and shorter operative times. Although patients in the OP-1 Putty group demonstrated an early propensity for formation of anti-OP-1 antibodies, this resolved completely in all patients with no clinical sequelae.

Conclusion. OP-1 Putty is a safe and effective alternative to autograft in the setting of uninstrumented posterolateral spinal arthrodesis performed for degenerative spondylolisthesis and symptomatic spinal stenosis.

Key words: spinal fusion, bone morphogenetic protein, spondylolisthesis. **Spine 2008;33:2850–2862**

Posterolateral spinal arthrodesis is commonly used for the treatment of patients with symptomatic degenerative spondylolisthesis unresponsive to nonoperative treatment. However, failure of fusion remains a common complication after surgery.[1,2] In addition to the lack of successful arthrodesis, donor site morbidity related to the bone graft harvest continues to present a problem affecting as many as 25% of patients after traditional spinal fusion using autogenous iliac crest bone graft.[3–6] Therefore, a plethora of bone graft extenders and alternatives have been developed in an attempt to improve the rates of healing and avoid the complications of autograft harvest.[7,8] The discovery of osteogenic proteins by Urist in the mid-1960s ushered in a new era of molecular biology in bone formation and healing.[9] This family of proteins has been subsequently named bone morphogenetic proteins (BMPs), and many members of this family have been isolated and characterized. BMPs exert their action by recruiting and stimulating pluripotent mesenchymal cells along an osteoblastic lineage resulting in the formation of bone.[10] Because of the powerful osteogenic potential of these proteins, they have been studied with considerable interest as a possible replacement or augmentation for autograft bone in the setting of spinal fusion. Several BMP preparations have been studied in preclinical and clinical trials for spinal applications.[11–16]

From the *Department of Orthopaedic Surgery, Thomas Jefferson University and The Rothman Institute, Philadelphia, PA; †Commonwealth Orthopaedics and Rehabilitation, Fairfax, VA; ‡Department of Radiology and Orthopedic Surgery, Yale University, New Haven, CT; §Department of Orthopaedic Surgery, William Beaumont Hospital, Royal Oak, MI; ¶Stryker Biotech, Hopkinton, MA; ‖Toronto Western Hospital, Toronto, Ontario; and **Denver Spine Center, Denver, CO.
Acknowledgment date: April 9, 2008. First revision date: June 5, 2008. Second revision date: July 23, 2008. Acceptance date: July 28, 2008.
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.
Corporate/Industry funds were not 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.*, honoraria, gifts, consultancies.
Address correspondence and reprint requests to Alexander R. Vaccaro, MD, Department of Orthopaedic Surgery and The Rothman Institute, 925 Chestnut Street, 5th Floor, Philadelphia, PA 19107; E-mail: alexvaccaro3@aol.com

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

Osteogenic Protein-1 (OP-1), also called recombinant human BMP-7 (rhBMP-7), is one such protein. OP-1 is a member of the TGF-β superfamily, and, like other members in this family, can induce the formation of bone when implanted in ectopic locations. Implants containing OP-1 and collagen matrix have been shown to be osteoinductive, osteoconductive, and to speed the rate of bone healing and to improve the performance of autograft in animals.[17–20] The human OP-1 gene has been cloned and introduced into a commercial cell line, facilitating the production of large quantities of recombinant human OP-1 (rhOP-1). OP-1, with various carrier preparations, has been studied in a number of animal models of spinal fusion.[12,13,21–23] The available human data involving BMPs suggest that these molecules are associated with a low risk of protein-related complications when used to promote bone healing or spinal fusion. These complications, although not recognized to date after the administration of OP-1, can consist of hypersensitivity to the administration of the protein, autoimmune reactions, or loss of efficacy of the protein at the intended target resulting from immune complex formation.

Several human pilot studies involving the use of OP-1 as both an adjunct to and a replacement for autograft in posterolateral spinal fusion studies have been performed to date. Fehlings and coworkers have reported that OP-1 can be used safely to achieve successful fusions in patients at higher risk for pseudarthrosis.[24,25] Conditions placing patients at higher risk over the general population after lumbar arthrodesis include nicotine usage, previous irradiation, administered chemotherapy, and continuous postoperative use of nonsteroidal anti-inflammatories. Vaccaro *et al* reported 1-year, 2-year, and minimum 4-year results of a prospective randomized, controlled, multicenter clinical pilot study comparing autograft *versus* OP-1 alone in the setting of uninstrumented posterolateral arthrodesis for degenerative spondylolisthesis.[26–28] These results consistently indicated the safety and efficacy of OP-1 and its comparability with autograft. At each time point, the groups treated with OP-1 demonstrated higher fusion rates, higher rates of clinical success (20% increase in the Oswestry scores), and no incidents of local or systemic toxicity, ectopic bone formation, or other adverse events related to the use of OP-1 Putty.

The purpose of this pivotal study was to establish the clinical and radiographic noninferiority of OP-1 Putty as a replacement for autograft bone when performing uninstrumented posterolateral spinal arthrodesis in a randomized controlled population of patients with symptomatic lumbar spinal stenosis and degenerative spondylolisthesis.

■ Materials and Methods

Study Design

This study was approved as an Investigational Device Exception study by the Food and Drug Administration and by the institutional review boards of the participating institutions. The design was a controlled, open-label (with blinded radiographic assessment), randomized, prospective, multicenter trial in which patients underwent single-level uninstrumented posterolateral lumbar arthrodesis for degenerative spondylolisthesis and spinal stenosis. The primary goal of the study was to demonstrate the safety and efficacy of OP-1 Putty and to demonstrate noninferiority *versus* the autograft control. The study was performed at 24 centers. After obtaining informed consent, patients were randomized to treatment with either OP-1 Putty or a control arm in which autogenous bone graft from the iliac crest (autograft) was used. A total of 335 patients were enrolled and randomized, of which 295 were treated. There was an attrition of 40 patients from the "intent-to-treat" population; 20 patients from the autograft group either refused the autograft part of the procedure or did not qualify after randomization based on the inclusion/exclusion criteria and 20 patients in the OP-1 group who were from the OP-1 Putty group either voluntarily withdrew from the study or were disqualified based on the inclusion/exclusion criteria. A total of 208 patients received OP-1 Putty and 87 received autograft. After surgery, patients were evaluated clinically and radiographically at 6 weeks, and at 3, 6, 9, 12, 24, and at a minimum of 36 months. Clinical assessments consisted of an evaluation of subjective pain and function using the Oswestry Low Back Pain Disability (ODI) questionnaire, the Visual Analog Scale (VAS), neurologic evaluation, and functional outcome assessment via completion of the Short-Form 36 (SF-36) outcomes survey. Imaging consisted of anteroposterior (AP), lateral, and flexion-extension radiographs. After the 24-month time point, patients were recruited to participate in the 36+ month assessment. At the latest follow-up at 36+ months, 202 of the original protocol patients (144 patients in the OP-1 Putty group and 58 patients in the autograft group) were evaluated with flexion-extension radiographs and helical CT scans with multiplanar reformatted imaging and three-dimensional (3-D) reconstructions. In addition, clinical assessments consisting of physical examination, SF-36 forms, and ODI questionnaires were repeated. Updated retreatment and serious adverse events (SAE) data were compiled through 36+ months.

Fusion Materials

A single package of OP-1 Putty implant consists of 3.5 mg of rhOP-1 formulated with 1 g of Type 1 bovine-derived collagen and 230 mg of carboxymethylcellulose. This powdered mixture was reconstituted at the time of surgery by the addition of saline to achieve a final implant concentration of rhOP-1 protein of 0.875 mg/mL. One package of implant was used per side, so that each patient received a total dose of 7 mg of rhOP-1 protein. No autogenous bone was used for the fusion in those patients randomized to receive the OP-1 Putty implants. Patients who were randomized to the autograft group were treated with corticocancellous bone harvested from the posterior iliac crest. No local bone graft was used for the fusion procedures. In both groups, the implanted fusion material was placed between the decorticated transverse processes and on the lateral border of the facets on both sides of the listhetic segment (*e.g.*, for a L4–L5 spondylolisthesis, the fusion material was used to bridge the space between the decorticated L4 and L5 transverse processes).

Inclusion and Exclusion Criteria

All study patients had Grade I or II degenerative spondylolisthesis of the L3–L4, L4–L5, or L5–S1 segments with coexistent spinal stenosis as confirmed by history, physical examination, and imaging, including AP and lateral plain radiography, flex-

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

ion-extension radiographs, and MRI or postmyelographic CT. Clinically, the patients presented with symptoms of neurogenic claudication. All the patients were skeletally mature, and none had undergone previous lumbar surgery. All patients had failed at least 6 months of nonoperative treatment, including physical therapy, lumbar epidural injections, anti-inflammatory medications, and activity modifications for their spinal symptoms. Exclusion criteria involved a spondylolisthesis of greater than Grade II, nondegenerative spondylolisthesis of any grade, spinal instability on flexion-extension radiographs measuring >50% translation of the vertebral body or >20° of angular motion, active spinal or systemic infection, systemic disease precluding participation (*e.g.*, neuropathy), current nicotine use, a history of smoking, morbid obesity, or a known sensitivity to collagen. Women of child-bearing potential who had not had a hysterectomy were also excluded.

Randomization and Demographics

Patients were randomized in a 2:1 ratio to receive OP-1 Putty or iliac crest autograft for the spinal arthrodesis aspect of the procedure. Randomization was performed after enrollment but before surgery using a computerized algorithm (SAS using the PLAN procedure). Patients and physicians became aware of the treatment assignment at the time of the randomization and before surgery so the study was unblinded; however, radiographic assessments of fusion and determination of neurologic success were performed by independent assessors in a blinded manner.

Surgical and Postoperative Protocol

All patients received general anesthesia and prophylactic antibiotics. A posterior midline exposure was performed and carried out to the tips of the transverse processes of the listhetic segment. A bilateral laminectomy and bilateral medial facetectomies were performed to decompress the neural elements. The transverse processes of the levels cephalad and caudad to the slip were decorticated to expose the marrow elements of the bone. The lateral border of the facets and the pars interarticularis were also decorticated. The fusion material (either 3.5 mg OP-1 Putty implant per side or half of the autograft bone graft per side) was placed in the intertransverse region to bridge the space between the decorticated transverse processes. Although a standardized technique was used to harvest corticocancellous bone from the posterior iliac crest, no formal method for quantification of the volume of autograft bone was used in the protocol. No irrigation was performed after placement of the fusion material.

Postoperative Management

Each patient was fitted with a lumbosacral orthosis of choice and instructed to wear the brace when out of bed for 3 months. Early ambulation was encouraged on the first day after surgery. Formal organized physical therapy emphasizing active exercises was begun 6 to 8 weeks after surgery. Each patient was scheduled for follow-up visits with their surgeon at 6-week, and 3, 6, 9, 12, 24-month time points after surgery. At each visit, a clinical neurologic and radiographic assessment was performed, including AP, lateral and flexion-extension plain radiographs (at the 3-month follow-up and later). Oswestry Disability Index (ODI), and Medical Outcomes Study 36-Item Short Form Health Survey (SF-36) were administered. VAS scores were assessed at the 12 and 24-month visits. At 36+ months, all available patients were brought back for clinical and radiographic reassessment. The criteria for recruitment of patients to the 36+ month group were that the patient had to be alive and had not been previously categorized as a retreatment failure earlier in the study. The repeat clinical assessments included ODI assessments, VAS scores, neurologic testing, retreatment analysis (*i.e.*, revision, removal, supplemental fixation, or reoperation) at the original treated level, and compilation of SAEs.

Radiographic Assessments

AP and flexion-extension radiographs through 24 months were interpreted by 2 radiologists blinded to the treatment group. A third, blinded radiologist evaluated those radiographs in which the 2 radiologists were in disagreement. Radiographs were used for the identification of the treatment level, assessment for bridging trabecular bone between the transverse processes, and identification of angulation and translation. The presence of new bone formation bridging across the transverse processes, angulation ≤5°, and ≤3 mm of translation were all required to meet the standard of radiographic fusion. The postoperative flexion-extension radiographs were performed as part of the assessment of radiographic fusion only. They were not performed to allow a comparison with preoperative flexion-extension values. Because the performed surgery involved a decompressive laminectomy and partial medial facetectomy of a degenerative spondylolisthesis, the postoperative condition of the spine before the development of fusion was considered to be unstable and, therefore, unsuitable for comparison with the preoperative condition. At the 36+ month interval, helical CT scans and flexion-extension plain films were taken and assessed using a prospective, multireviewer, blinded radiographic assessment protocol designed to minimize bias. The CT scans were performed using a standardized imaging algorithm and protocol to assess the presence of new bone formation and the presence of bridging bone across the transverse processes. The CT scans were also used to determine the location of bone formation (medial *vs.* lateral, in reference to the transverse process and pars interarticularis). Medial bone formation was determined to be across the pars interarticularis or the medial one-third of the transverse process region, whereas lateral bone formation was defined as bone formation extending across the lateral two-thirds of the transverse process region. Both the 36+ month CT scans and flexion-extension plain radiographs were interpreted by 2 primary spine surgeon readers not associated with the clinical trial and blinded to the treatment arm. When the 2 primary readers did not agree, a third reader was used to adjudicate the results and the majority assessment was used.

Immunologic Assessments

Enzyme-linked immunosorbent assays (ELISA) were performed to detect the presence of anti-OP-1 antibodies in all samples. ELISA methods were validated to detect human anti-human OP-1 antibodies with IgG, IgM, and IgE isotypes. The ELISA cutoff point for this study was statistically based and reflects a false positive rate of 5%, as recommended by Mire-Sluis *et al.*[29] Positive samples in screening ELISA were considered potentially positive for anti-OP-1 antibodies and tested in a validated confirmatory competition ELISA. Positive samples in the competition ELISA were further evaluated in a titer ELISA to quantify the level of anti-OP-1 antibodies in the sample. The results of this assay are reported as a log titer, which corresponds to the log of the lowest dilution of the sample that yields a positive result.

Samples found to be positive in the titer ELISA were further analyzed to determine whether antibodies to OP-1 had the

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

ability to neutralize its activity *in vitro*. Samples were initially tested in a luciferase reporter-based primary neutralizing antibody assay (nab). The presence or absence of antibodies (both anti-OP-1 antibodies and anti-OP-1 neutralizing antibodies) was determined following blood draw and centrifugation using ELISA analysis (Genetics Institute, Cambridge, MA). All patients who were antibody-positive at 24 months had repeat serum samples obtained at the 36+ month visit.

Primary Outcome Assessments

Safety and Adverse Outcome Reporting. The safety of the investigational product was evaluated by comparing the nature and frequency of adverse events in each of the 2 treatment groups. Adverse events included all minor and major medical events for which the patient sought medical attention regardless of the nature of the event or its severity. An adverse event was defined as any clinically adverse sign, symptom, syndrome, or illness that occurred or worsened during the operative or postoperative period of the trial, regardless of causality. All reoperations (revisions or supplemental fixations) over the study period were recorded. Reoperations performed to promote fusion at the treated level were deemed failures. Laboratory testing for immunologic, hematologic, and biochemical evaluation was performed before surgery (baseline), at 6 weeks, and at 3, 6, 12, and 24 months.

Primary End Points

The primary end points for the study were evaluated at 24 and at 36+ months. The primary end point at 24 months was designed for FDA submission evaluating the safety and efficacy of OP-1 Putty as a replacement for autologous iliac crest in the setting of a posterolateral fusion for degenerative spondylolisthesis. Primary Overall Success at 24 months was defined as a composite measure that required a 20% improvement in ODI, absence of treatment-emergent SAEs related to the treatment device, absence of a decrease in neurologic status (assessing muscle strength, reflexes, sensation, and straight leg raise), and radiographic fusion success. Radiographic fusion success was also a composite measure, requiring the presence of bridging bone as assessed on AP radiographs, angular motion ≤5°, and translational movement ≤3 mm as assessed by flexion-extension radiographs. The primary outcome assessment for the study at 36+ months, Modified Overall Success, was also defined as a composite measure requiring success on each of the following components: improvement of at least 20% in the ODI from baseline, absence of treatment-emergent SAEs related to the treatment device, absence of a decrease in neurologic status (assessing muscle strength, reflexes, sensation, and straight leg raise) at 24 months, and presence of new bone formation by CT scan, angulation of ≤5° and translational movement of ≤3 mm on flexion/extension radiographs, and absence of retreatment intended to promote fusion at 36+ months.

Data Analysis and Statistics

A power analysis performed before the initiation of the study demonstrated, using an alpha level of 0.05 and a power of 80%, that 270 treated subjects (180 OP-1 Putty, 90 autograft) were needed for the study. The number of treated patients in this trial was based on hypothesized overall success rates of 53% for the OP-1 Putty group in comparison with 47% for the autograft group based on data from a pilot study conducted on a similar population of patients with a similar endpoint. The maximum allowable difference between the treatment groups that could be used to conclude that OP-1 Putty was not inferior to autograft was variable.

Continuous variables were summarized using descriptive statistics (mean, median, standard deviation, minimum, maximum). Categorical variables were summarized using frequencies and percentages. Inferential tests were performed at the 5% level of significance.

The primary efficacy end points were the 24-month and the 36+ month overall success rates. The 36+ month rate of overall success included the 24-month overall success rate data with radiographic and retreatment (need for revision surgery at the index surgical level) data at 36+ months for the intent-to-treat population with missing data imputed using a multiple imputation technique. The percentage of successes (and standard error) was based on estimates of the treatment effect adjusted for covariates in logistic regression and on variance estimates obtained from multiple imputations. Secondary efficacy end points included analyses of overall success stratified by center size, age category, and gender.

For imputed modified overall success at 24 months (with radiographic and retreatment data at 36+ months), a one-sided two-sample asymptotic test for noninferiority was used. For both the primary efficacy analyses of success, the 95% upper confidence bound was generated corresponding to the difference in success rates (autograft minus OP-1 Putty) in the 2 treatment groups.

For adverse events, each SOC and each preferred term reported by ≥5% of patients in either treatment arm were tested for treatment differences using Fisher exact test. For neurologic status, χ^2 or Fisher exact test was used to test the difference between treatments groups and McNemar's test was used to test the shifts in status within treatment group.

Results

Demographic Information

Demographic and baseline data for the patients enrolled in the study are presented in Table 1. Overall mean age at baseline was 68 years (range 36–84 years). There were no significant differences between the OP-1 and autograft groups with respect to age, gender, weight, height, level treated, preoperative ODI, preoperative translation, or diagnosis.

At the 36+ month assessment time point, 80% (80.5%) of eligible patients (79.7% of autograft group and 80.8% of the OP-1 Putty group) returned or had died before study follow-up and were, therefore, accounted for in the long-term evaluation. All key demographic characteristics and 24-month outcome variables of the patients who participated in the long-term evaluation compared to those eligible to participate were similar and not statistically different (Table 2).

Surgical Indications and Prior Treatments

The indications for surgery are summarized in Tables 1 and 2. At baseline, all the patients carried a diagnosis of degenerative lumbar spondylolisthesis with spinal stenosis. Of these 272 of 293 (92.8%) had Grade I spondylolisthesis by the Meyerding classification,[30] 10 of 293 (3.4%) had Grade II, and 11 of 293 (3.8%) had spondylolisthesis that could not be distinguished between

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

Table 1. Demographics and Baseline Characteristics (Modified ITT)

Parameter	Statistic	Mean	OP-1 Putty	Autograft	*P*
Age (yrs)					
Mean		68	68	69	0.129
Median		69	68	71	
Std. Dev.		9.4	9.8	8.3	
Sex					
Male	N (%)	97 (33.1)	71 (34.3)	26 (30.2)	0.501
Female	N (%)	196 (66.9)	136 (65.7)	60 (69.8)	
Level fused					
L3–L4	n (%)	31 (10.6)	21 (10.1)	10 (11.6)	0.76
L4–L5	n (%)	252 (86.0)	178 (86.0)	74 (86.0)	
L5–S1	n (%)	10 (3.4)	8 (3.9)	2 (2.3)	
ODI	N	293	207	86	
Mean		48.8	48.8	48.8	0.998
Median		48	48.9	48	
Std. Dev.		12.19	11.6	13.59	
Angular motion (degrees)	N	271	195	76	
Mean		4.1	3.9	4.7	0.086
Median		3.1	2.8	4.2	
Std. Dev.		3.36	3.4	3.2	
Translational movement (mm)	N	268	193	75	
Mean		1.7	1.7	1.6	0.802
Median		1.4	1.4	1.1	
Std. Dev.		1.45	1.44	1.49	
Diagnosis of degenerative lumbar spondylolisthesis with spinal stenosis	n (%)	293 (100.0)	207 (100.0)	86 (100.0)	—
Grade 1	n (%)	272 (92.8)	193 (93.2)	79 (91.9)	
Grade 2	n (%)	10 (3.4)	8 (3.9)	2 (2.3)	
Unable to distinguish between Grade 1/2	n (%)	11 (3.8)	6 (2.9)	5 (5.8)	

Grade I and Grade II. Two hundred fifty-two of 293 (86.0%) patients had disease at the L4–L5 level, 31 of 293 (10.6%) patients had disease at the L3–L4 level, and 10 of 293 (3.4%) patients had disease at the L5–S1 level.

All patients had failed at least 6 months of nonoperative treatment, including physical therapy, lumbar epidural injections, anti-inflammatory medications, and activity modifications for their spinal symptoms.

Table 2. Demographic and Baseline Characteristics for 36+ Months Eligible Patients

		OP-1 Putty		Autograft	
Parameter	Statistic	24 mo	36+ mo	24 mo	36+ mo
Age (yrs)	N	183	144	74	58
Mean		67.6	66.8	69.3	68.7
Median		69	67	71	70
Std. Dev.		9.54	9.25	8.72	8.66
Sex					
Male	N (%)	64 (35.0)	50 (34.7)	19 (25.7)	16 (27.6)
Female	N (%)	119 (65.0)	94 (65.3)	55 (74.3)	42 (72.4)
Level fused					
L3–L4	n (%)	19 (10.4)	17 (11.8)	10 (13.5)	9 (15.5)
L4–L5	n (%)	156 (85.2)	124 (86.1)	62 (83.8)	48 (82.8)
L5–S1	n (%)	8 (4.4)	3 (2.1)	2 (2.7)	1 (1.7)
ODI	n	183	144	74	58
Mean		48.5	48.2	50.1	50.7
Median		48.9	48.9	48	48
Std. Dev.		11.11	10.74	13.48	12.47
Angular motion (degrees)	n	174	138	66	51
Mean		4	4.1	4.7	4.3
Median		2.8	2.9	4.1	3.6
Std. Dev.		3.42	3.53	3.24	3.03
Translational movement (mm)	n	171	136	65	51
Mean		1.7	1.8	1.6	1.5
Median		1.4	1.5	1	0.8
Std. Dev.		1.48	1.55	1.52	1.42
Diagnosis of degenerative lumbar spondylolisthesis with spinal stenosis	n (%)	183 (100)	144 (100)	74 (100)	58 (100)
Grade 1	n (%)	169 (92.3)	135 (93.8)	68 (91.9)	54 (93.1)
Grade 2	n (%)	8 (4.4)	5 (3.5)	2 (2.7)	2 (3.4)
Unable to distinguish between Grade 1/2	n (%)	6 (3.3)	4 (2.8)	4 (5.4)	2 (3.4)
Prior overall success at 24 mo	%	42.9	43.8	57.6	60

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

Table 3. Overall Success at 24 Months (MITT)

Parameter	OP-1 Putty	Autograft	*P* for Noninferiority
Overall success	38.7%	49.4%	0.33

Of the 202 patients available at 36+ months, 169 of 183 (92.3%) had Grade I spondylolisthesis, 8 of 183 (4.4%) had Grade II spondylolisthesis, and 6 of 183 (3.3%) had spondylolisthesis that could not be distinguished between Grade I and Grade II. One hundred fifty-six of 183 (85.2%) patients had disease at the L4–L5 level, 19 of 183(10.4%) patients had disease at the L3–L4 level, and 8 of 183 (4.4%) patients had disease at the L5–S1 level.

36+ Month Follow-up Evaluation

Of the 257 patients eligible for 36+ month follow-up, 80.5% (202/257) were available. These consisted of 144 of the original 208 OP-1 patients (69%) and 58 of the original 87 (67%) autograft patients. Of the 55 of 257 who did not have 36+ month follow-up, 5 patients had died, 23 refused to participate, 15 could not be located, 3 were unable to participate because of the unavailability of the participating site, and 9 could not participate for other reasons. Therefore, 80.5% (202/257) of the eligible patients (79.7% of the autograft group, and 80.8% of the OP-1 Putty Group) at 24 months were accounted for at the time of final follow-up (36+ months). The mean time to final follow-up was 4.4 years (range 3.68–5.46, SD 0.4) for all enrolled patients, 4.38 years (range 3.68–5.42, SD 0.4) for the OP-1 Putty group, and 4.47 years (range 3.76–5.46, SD 0.4) for the autograft group. There were no significant difference in times to follow-up ($P = 0.143$). Although a small percentage of the total eligible population was lost to follow-up at the 36+ month interval, the reasons for nonparticipation in 36+ month extension part of the study were equally distributed across the study groups.

Primary Outcome

The Overall Success outcome composite end point at 24 months revealed that statistical equivalence was not achieved between the 2 groups at 24 months (38.7% for the OP-1 Putty group and 49.4% for the Autograft group, $P = 0.33$) (Table 3). Among the subcomponents of Overall Success, there was a statistically significant difference between the groups in terms of the presence of bridging bone as assessed by plain radiographs (61.7% for the OP-1 Putty group and 83.1% for the Autograft group, $P < 0.001$). There were no other statistically significant differences among the clinical or radiographic subcomponents of Overall Success (Table 4).

Table 4. Subcomponents of Overall Success at 24 Months

Parameter	OP-1 Putty	Autograft	*P* for Difference
Components of overall radiographic success			
Presence of bridging bone by plain film	61.7%	83.1%	<0.001
Angulation ≤5° on flexion/extension films	73.3%	75.6%	0.684
Translation ≤3 mm on flexion/extension films	87.7%	87.8%	0.978
ODI success	74.5%	75.7%	0.839
Absence of retreatment	92.3%	88.6%	0.347
Absence of serious treatment-related AEs	85.6%	84.7%	0.863
Neurological success	92.1%	84.1%	0.057

Table 5. Modified Overall Success at 36+ Months

Parameter	OP-1 Putty	Autograft	*P* for Noninferiority
Overall success	47.2%	46.8%	0.025

The Modified Overall Success outcome end point at 36+ months revealed no difference and statistical comparability between the 2 study groups (47.2% for the OP-1 Putty group and 46.8% for the Autograft group, $P = 0.025$) (Table 5). Furthermore, there were no statistically significant differences among the subcomponents of Modified Overall Success at 36+ months (Table 6).

Imaging Findings

At 24 months, 73.3% of OP-1 subjects and 75.6% of autograft subjects had ≤ 5° of angular motion ($P = 0.684$) and at 36+ months 69.3% of OP-1 subjects, and 68.4% of autograft subjects had ≤5° of angular motion ($P = 1.0$). At 24 months, 87.7% of OP-1 subjects and 87.8% of autograft subjects had ≤3 mm translation ($P = 0.978$). At 36+ months, 75.7% of OP-1 subjects and 75.4% of autograft subjects had ≤3 mm translation ($P = 1.0$). There were no statistical differences between the study groups in terms of angular or translational motion at either time point.

CT scans were obtained on 196 of 202 (97%) patients available at 36+ months: 143 from the OP-1 Putty group and 53 from the autograft group. One hundred seven of 143 (74.8%) of the OP-1 Putty patients and 41 of 53 (77.4%) of the autograft patients had presence of new bone on CT scan. The results were clinically comparable and not statistically significantly different ($P = 0.852$).

Table 6. Subcomponents of Modified Overall Success at 36+ Months

Parameter	OP-1 Putty	Autograft	*P*
Presence of Bone on CT Scan (36+ months)	74.8%	77.4%	0.852
Angulation ≤5° on flexion/extension films (36+ months)	69.3%	68.4%	1
Translation ≤3 mm on flexion/extension films (36+ months)	75.7%	75.4%	1
ODI success (24 mo)	74.5%	75.7%	0.839
Absence of retreatment (36 mo)	87.0%	83.3%	0.529
Absence of serious treatment-related adverse events (24 mo)	85.6%	84.7%	0.863
Neurologic success (24 mo)	92.1%	84.1%	0.057

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

Table 7. Presence of Bone Assessed *via* CT Scan for Patients Without Bone Formation *via* Plain Films at 24 Months (mITT)

Presence of Bone Assessed *via* CT Scan	OP-1 Putty	Autograft
Present	27/38 (71.0%)	5/6 (83.3%)
Present medial	22/27 (81.5%)	3/5 (60.0%)
Present lateral/transverse	5/27 (18.5%)	2/5 (40.0%)
Absent	11/38 (29.0%)	1/6 (16.7%)

Analysis of the OP-1 Putty patients who had previously been assigned as failures because of lack of bone formation (38 patients) based on the 24-month plain radiographs and had undergone CT scans at 36+ months demonstrated that 27 of 38 (71%) exhibited bone formation on CT scan. Furthermore, of these 27 patients who did form bone, 22 of 27 (81.5%) were found to have bone formation that was classified as medial, and 5 of 27 (18.5%) had bone formation that was classified as lateral (Table 7).

Although stability of the fusion was not different between the treatment groups, the presence of bridging bone across the intertransverse process region was dissimilar in the 36+ month CT data. Bridging bone was detected in 56% of patients in the OP-1 Putty group and 83% ($P = 0.001$) of patients in the autograft group. Detection of bridging bone was more difficult medially as shown by a higher disagreement rate between the readers' assessment of bridging bone with OP-1 Putty (29%) compared with autograft (8%) ($P = 0.0124$).

Device-Related Serious Adverse Events and Retreatment Failures

Table 8 presents the success rate based on the absence of treatment-related SAEs categorized as related to the device. At 24 months, the OP-1 Putty group exhibited a higher proportion of patients free from treatment-related SAEs than did the autograft group (85.6% for OP-1 Putty and 84.7% for autograft, $P = 0.863$). At 36+ months the OP-1 Putty group again experienced a higher proportion of patients free from treatment-related SAEs (79.5% for OP-1 Putty and 73.5% for autograft, $P = 0.387$). These differences did not reach statistical significance.

Retreatment Failures

The OP-1 Putty group demonstrated a higher proportion of patients who were free from retreatment failures at 24 months (179/194 patients, or 92.3% for OP-1 Putty and 62/70 patients, or 88.6% for autograft, $P = 0.347$) and at 36+ months (141/162 patients, or 87.0% for OP-1 Putty and 55/66 patients, or 83.3% for autograft, $P = 0.529$). These values were statistically not different.

Table 8. Success Rates by Absence of Treatment-Related Serious Adverse Events (SAEs) at 24 and 36+ Months

	OP-1 Putty		Autograft		
Time Point	No. Patient	No. (%) Successes	No. Patient	No. (%) Successes	*P*
24 mo	194	166 (85.6)	72	61 (84.7)	0.863
36+ mo	166	132 (79.5)	68	50 (73.5)	0.387

There were 32 total retreatments (*i.e.*, failures for the absence of retreatment criteria): 21 in the OP-1 Putty group (17 reported at 24 months and 4 reported at 24–36+ months) and 11 in the autograft group (10 reported at 24 months and 1 reported at 24–36+ months). Retreatments occurring over time for both treatment groups are illustrated in Table 9. In both treatment groups, the majority of retreatment events occurred in the interval between the immediate postoperative period and the 24-month interval, with the balance of events occurring by the 48-month interval, and no events occurring at or after the 60-month interval.

At the 24-month follow-up point, 21 of 257 (8.2%) of the OP-1 Putty patients and 11 of 87 (13%) of the autograft patients had undergone further surgery to promote fusion at the index level. At the time of latest follow-up (>36 months), an additional 3 of 144 (2.1%) OP-1 Putty patients and 3 of 58 (5.2%) autograft patients had undergone further surgery for retreatment failure. These rates were not statistically different ($P = 0.242$).

Secondary Outcomes

Oswestry Disability Index. Tables 10 and 11 present the success rates for the study groups at 24 and 36+ months as measured by a 20% improvement in the ODI. At 24 months, 74.5% of OP-1 subjects and 75.7% of autograft subjects had a ≥20% improvement from baseline in ODI. At 36+ months, 68.6% of OP-1 subjects and 77.3% of autograft subjects had a ≥20% improvement from baseline in ODI. There were no statistical differences between the groups at either time point ($P = 0.839$ at 24 months, $P = 0.201$ at 36+ months). The mean percent improvements from baseline 24 months (54.0% for OP-1 Putty and 54.5% for autograft) and 36+ months (52.0% for OP-1 Putty and 54.4% for autograft) were similar and not statistically different between treatment groups.

Because the 20% ODI improvement from baseline is an arbitrary cut point for determining clinical improvement, additional analyses were conducted to compare the proportions of patients in each treatment group achieving more robust levels of improvement that should be more clinically meaningful to both physicians and patients. The number of patients in each treatment group achieving improvements over baseline of 100%, ≥80%, ≥50%, ≥30%, and ≥20% at both 24 months and 36+ months was evaluated (Figure 1). These results indicate that although the OP-1 Putty group had slightly lower proportions of patients who achieved ODI success in the ≥20% and ≥30% improvement in ODI categories (differences not statistically significant), the OP-1 Putty group had higher proportions of patients achieving

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

Table 9. Retreatment Failures by Time Interval (Safety Population)

Treatment Group	Operative (<28 d)	24 mo (28–1035 d)	36 mo (1036–1401 d)	48 mo (1402–1766 d)	60 mo (1767–2131 d)	>60 mo (>2131 d)	Totals
OP-1 Putty	0	17	3	1	0	0	21
Autograft	0	10	0	1	0	0	11

≥50%, ≥80%, and 100% improvements at both the 24 month and 36+ month intervals (differences not statistically significant).

Neurologic Success

The patient was considered an overall neurologic success in the absence of a decrease in neurologic status unless attributable to a concurrent medical condition or to the surgical procedure. Patients in the OP-1 Putty group had a higher neurologic success rate at 24 months (92.1% for OP-1 Putty and 84.1% for autograft, $P = 0.057$), although this difference was not statistically significant. Neurologic success was similar for both groups at 36+ months, and the difference between treatment groups was not statistically significant (84.4% for OP-1 Putty and 80.0% for autograft, $P = 0.54$).

Visual Analog Scale and Short-Form 36

Patients in both the OP-1 Putty and Autograft Groups had significant decreases in pain over time noted on VAS at 24 months and at 36 months. There were no significant differences between the 2 groups in terms of VAS scores. By 6 weeks, patients in both groups demonstrated statistically significant improvements over baselines in SF-36 scores. There were no significant differences between group SF-36 scores at any point in the study (Figure 2).

Donor Site Pain After Autograft Harvest

VASs assessments of donor site pain in the autograft population demonstrated that at 12 months, 32 of 72 (44%) of autograft patients reported pain at the donor site, at 24 months 25 of 55 (45%) patients reported pain, and at 36 months 18 of 52 (35%) reported persistent mild/moderate pain. Donor site pain was persistent and decreased slowly over time, reported as a 2.1 on the VAS (scale of 1–10, 10 being most severe) at 6 weeks, 1.6 at 12 months, 1.2 at 24 months, and 1.1 at 36 months.

Surgery and Hospitalization Data

Mean operative time for the OP-1 Putty group was significantly shorter than the autograft group (144 minutes for the OP-1 Putty group and 164 minutes for the autograft group, $P = 0.006$). Mean operative blood loss was also significantly lower for the OP-1 Putty group than the autograft group (309 cc *vs.* 471 cc, $P = 0.00004$). There were no differences in the mean length of stay after surgery ($P = 0.529$).

Immunologic Results

Serum samples for OP-1 antibody testing for the study were performed immediately after surgery, and at 6 weeks, 3, 6, 12, and 24 months from 293 patients. One patient in the OP-1 group had died just after surgery and 1 patient in the autograft group had no postbaseline visit. In the 36+ month group, serum samples were analyzed at the time of latest follow-up for patients who had been positive for anti-OP-1 antibodies at the 24-month follow-up visit and for patients who had not completed the 24-month follow-up visit but had been antibody positive at their last recorded visit. There were 54 patients who underwent this testing (49 patients from the OP-1 Putty group, 5 from the autograft group) at 36+ months.

93.7% of patients receiving OP-1 Putty were antibody-positive at any time point *versus* 20.9% of the patients receiving autograft. In the OP-1 Putty group, 25.6% of patients became positive for anti-OP-1 neutralizing antibodies *versus* 1.2% of the autograft patients. The peak presence of neutralizing antibodies was observed between 6 weeks and 3 months. However, at both 24 and 36+ months no patients had neutralizing antibodies present.

Neutralizing Activity Status and Clinical Outcomes

No significant associations were observed between neutralizing activity status, clinical success, and safety parameters. Overall success of patients with neutralizing activity (36.4%) was not statistically different from the overall success of patients without detectable neutralizing activity (38.2%). When the overall success end point was broken down into the individual components of radiographic success, ODI success and absence of retreatment, no associations between clinical success and neutralizing activity were seen. Furthermore, there was no evidence of an increase in AEs, SAES, or immunologically-related AEs or SAEs at any time point in the neutralizing positive patients *versus* the neutralizing negative patients in the study.

Table 10. ODI Scores at 24 Months

	OP-1 Putty		Autograft			
Time Point	No. Patients	No. (%) Successes	No. Patients	No. (%) Successes	95% Upper Confidence Bound (%)	*P*
24 mo	192	143 (74.5)	70	53 (75.7)	11.2	0.839

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

Table 11. ODI Scores at 36+ Months (mITT Population)

	OP-1 Putty		Autograft		
Time Point/Population	No. Patients	No. (%) Successes	No. Patients	No. (%) Successes	*P*
36+ mo	159	109 (68.6)	66	51 (77.3)	0.201

■ Discussion

This study was a randomized controlled trial comparing 2 similar groups of patients with degenerative spondylolisthesis and spinal stenosis. The study population was reflective of the general population with degenerative spondylolisthesis, given the relatively higher number of women (approximately 2/3) and the mean age of 68 (range 36–84). The preoperative status of the patients in the respective study groups was similar, with no differences between groups in terms of demographic characteristics, disease status, the involved segment, motion or instability at the involved level, clinical status based on the ODI, previous treatments, or worker's compensation status.

This pivotal study was originally designed to report on patient outcomes as part of the 24-month randomized, prospective, multicenter trial conducted under an Investigational Device Exemption Study as permitted by the US FDA. In the original 24-month investigational study, patients in the OP-1 Putty group achieved clinical and functional radiographic improvements comparable to the autograft group measured along a composite measure of clinical success. Although there were no statistical differences between the OP-1 Putty group and the autograft group in terms of overall success by these composite end points, there was a statistical difference between the groups in terms of the presence of bridging bone on plain radiography. As assessed by plain films, the OP-1 Putty group demonstrated a significantly lower percentage of patients with presence of bridging trabecular bone. There were no differences seen in angulation or translation on flexion-extension films suggestive of a true difference in the fusion mass. A long-term follow-up was performed to see if clinical and radiographic results were maintained over time and to see if patients who seemed to demonstrate radiographic success at 2 years maintained success as reported in other clinical IDE fusion studies.[31] One interesting finding in this study was the presence of bone formation in the OP-1 Putty fusion group medially along the transverse processes and along the lateral border of the facet joints on the 9-month and 36+ month CT scans. These results suggest that plain films may be less than reliable in assessing fusion or bone formation with the present physical formulation of OP-1 Putty, as they are less sensitive when compared with CT in assessing bone formation along the lateral border of the facet joints.

The finding of bone formation medial to the transverse processes was unexpected, because it had been assumed that OP-1 Putty-directed new bone formation would occur as it does for autograft, laterally along the transverse processes. A probable explanation lies in the differences in the physical properties of the graft materials studied: OP-1 Putty is a compressible, moldable material (putty that does not harden), whereas autograft is not malleable (has a noncompressible physical structure). During the spinal fusion procedure used in the clinical study, the surgeon retracts the paraspinal muscles to lay down the OP-1 Putty or autograft material (Figures 3A, B). When the retractors are removed and the muscles are released, the OP-1 Putty product may be compressed medially (Figure 3C), leading to medialized bone formation not easily detected by plain radiographs. On plain radiographs, the medial location of the OP-1 putty may be obscured by the lateral border of the ver-

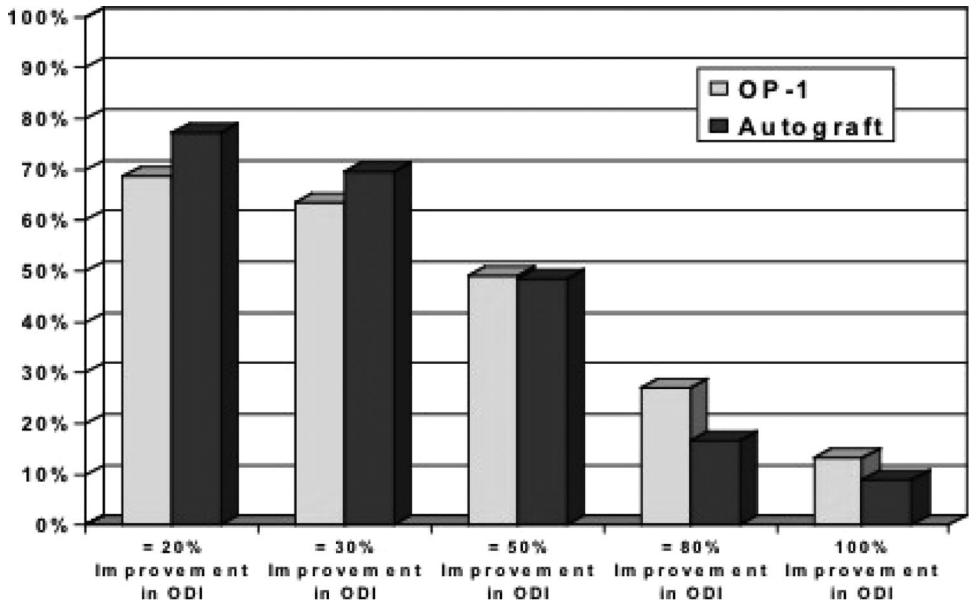


Figure 1. Proportions of patients with percentage improvements in ODI at 36+ months.

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

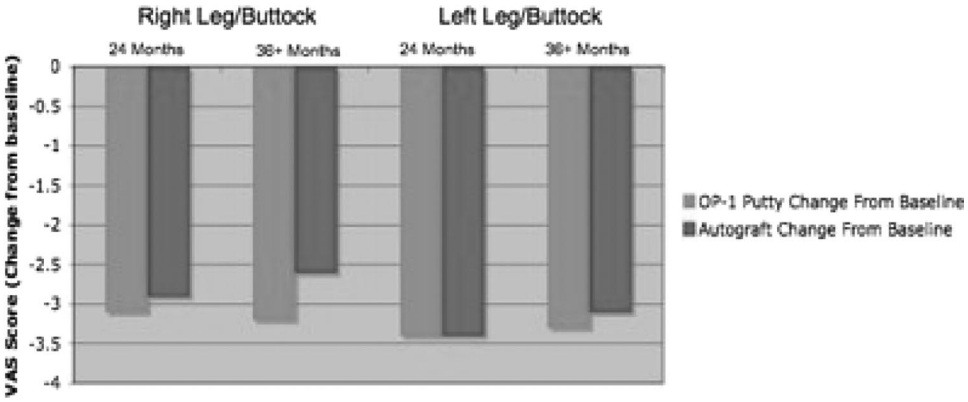


Figure 2. VAS scores at 24 and 36+ months.

tebral body, hypertrophied facet joints, and overlying bowel gas (Figure 4A), but may be better illustrated with the usage of CT scanning (Figure 4B). A similar phenomenon was reported by Boden and coworkers during a study of posterolateral intertransverse process fusion using a compressible collagen carrier and rhBMP-2 in a primate model.[32]

Therefore, to adequately investigate whether patients in the OP-1 Putty group experienced fusion rates comparable to autograft, a radiographic assessment tool more sensitive than the plain films used at 24 months was needed. As a result, prospective collection of additional radiographic and clinical data was conducted on all available study patients at the longer term follow-up interval of 36+ months. Patients received CT scans to assess for the presence of bone and repeat flexion/extension films to allow measurement of angulation and translation at the same time point. All key clinical outcome measures collected in the original study were also collected at the 36+ month interval.

Eighty-seven percent of the originally randomized 295 patients were eligible for review at 24 months. Of the 257 patients eligible for 36+ month follow-up, 202 of 257 (80%) were available. Patients who returned for follow-up at 36+ months did not differ significantly from the original population with regard to demographics, baseline disease characteristics, or key outcome variables at 24 months. The mean time to final follow-up for all enrolled patients was 4.4 years and there was no significant difference in times to follow-up between treatment groups.

At 36+ months, OP-1 Putty demonstrated statistical comparability to autograft with regard to the primary endpoint (Modified Overall Success), and each of the individual subcomponents. Although the 24-month radiographic data did not show comparability of OP-1 Putty and autograph in terms of rates of new bone formation by plain radiograph despite comparable results for angulation and translation, the 36+ month imaging data found no difference between groups in terms of the presence of bone on CT (74.8% for OP-1 Putty and 77.4% for autograft, $P = 0.85$). There are several possible reasons for these differences in findings. First, the original readers may have underestimated the presence of bone as they were not aware of the medial repositioning of OP-1 Putty. Second, CT scanning represents a more sensitive imaging modality for bone formation given the increase in spatial resolution. Finally, the CT scan data demonstrate that the OP-1 Putty group formed bone in a more medial location with greater frequency (Figures 5A, B).

Because of the medial fusion mass when OP-1 Putty is used, CT assessment of fusion may not accurately reflect the degree of bridging bone. Although some patients in the OP-1 Putty arm did show bridging bone between the transverse processes, the majority of the bone formation was more medial. CT imaging of the spine has improved greatly over the last decade with advancements in image acquisition. However, the radiologist must review an increased number of slices and images and visualize a complex 3-D model. Bone seen in the area of the transverse process most likely represents graft because degenerative

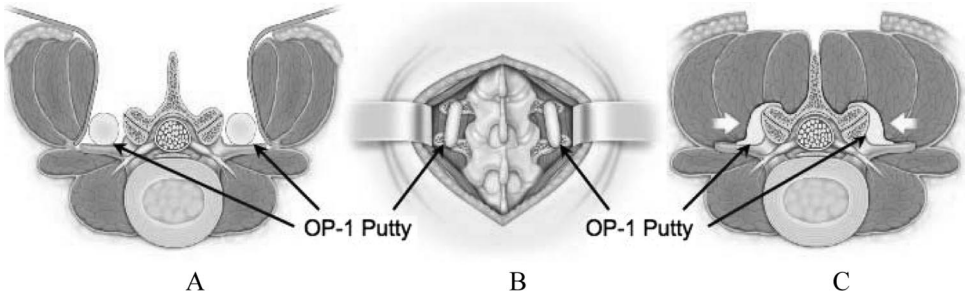


Figure 3. Implantation of OP-1 Putty and the paraspinal musculature. **A,** Axial illustration demonstrating retraction of paraspinal muscles and placement of OP-1 Putty across the decorticated transverse processes. **B,** Coronal view demonstrating OP-1 Putty placement across the intranverse region. **C,** After paraspinal muscle release, OP-1 Putty is compressed medial to the transverse processes.

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

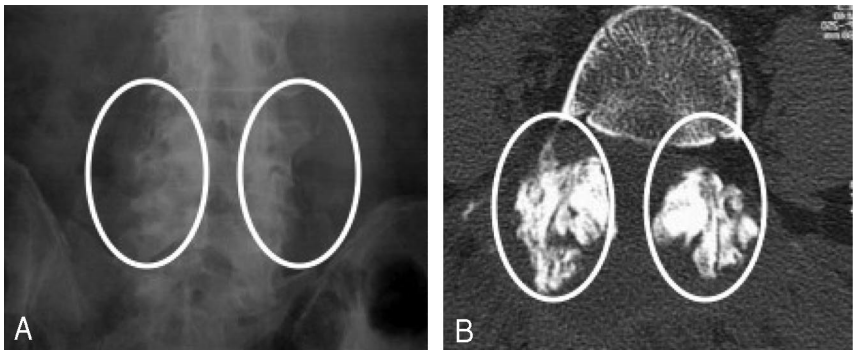


Figure 4. AP radiograph (**A**) taken at 24 months compared to axial CT (**B**) scan in a patient after receiving OP-1 Putty. While the radiograph (**A**) fails to illustrate bridging bone between the transverse processes, axial CT scan (**B**) demonstrates profuse bone formation more medially. The obtained plain radiograph was interpreted by blinded observers as having no bone formation at 24 months.

changes in the spine are not seen in that area. If the same amount of bone is seen more medially near the facet, it could easily be misinterpreted as an osteophyte rather than bone being generated from biologic material placed medially. Because bone that has formed medial to the transverse processes is more difficult to assess for bridging, one could easily underestimate the degree of fusion with OP-1 compared to autograft across transverse processes. A future prospect may lie in the use of multiplanar reformatted images and the ability to generate 3-D models (Figures 6A and 6), which may serve to provide more usable information on bony fusion.

With regard to fusion outcomes, it is interesting to compare bone formation rates in the 50% to 70% range relative to previously published studies on bone morphogenetic factors with claimed fusion success rates of 95% to 100%. This comparison highlights the strength of a prospective randomized study evaluating fusion success in an unstable (spondylolisthesis) degenerative model in the absence of instrumentation or an opaque carrier, which may be confused with new bone formation. In the seminal paper by Fischgrund *et al* on fusion success in the presence or absence of instrumentation, the authors noted a fusion success of only 45% in the absence of instrumentation.[1] Clearly, without confounding variables such as instrumentation or carriers that contain calcium or hydroxyapatite, fusion rates in degenerative disorders are expected to be in the 40% to 70% success rate range, depending on patient characteristics. This further supports the premise that OP-1 Putty in the stated dosage and with its compressible carrier is an adequate replacement for autologous iliac crest bone graft in fusion for degenerative spondylolisthesis. With use of instrumentation, the current gold standard for treatment, higher rates of fusion and overall success would be anticipated. All patients in this study were considered to be "unstable" before fusion because of the destabilization occurring secondary to the laminectomy and partial facetectomy. When evaluating the final radiographs, the absence of motion on flexion and extension views may be because of bone formation from either OP-1 Putty, autograft, or a stable fibrous nonunion. It is assumed that successful fusion occurs when there is sufficient bone formation to confer stability to the spine (by meeting the stringent translation and angulation criteria for successful fusion). This argument would be invalid if instrumentation had been placedbecause of the presence of hardware restricting motion.

Patients receiving OP-1 Putty as part of the arthrodesis surgery had significantly lower blood loss at the time of surgery and lower operative times. Although these numbers did not result in a lower rate of treatment-related serious adverse events, the decreased operative time, decreased exposure to anesthesia, and expected lower transfusion requirements are potential benefit to this elderly surgical population in terms of an expected quicker recovery and lower rate of transfusion-related complications, in addition to providing a possible eco-

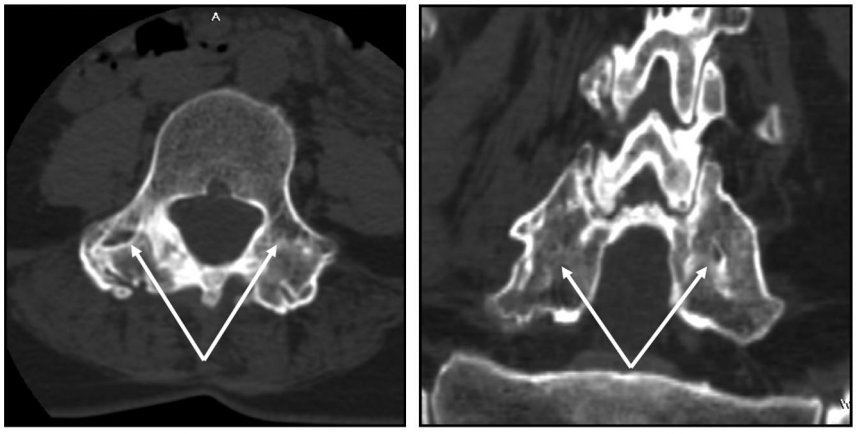


Figure 5. Axial CT (**A**) and coronal multiplanar reformatted image (**B**). Axial CT and multiplanar reformatted coronal images demonstrate a solid fusion mass in a patient receiving OP-1 Putty. Note the medial positioning of the fusion mass (white arrows).

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

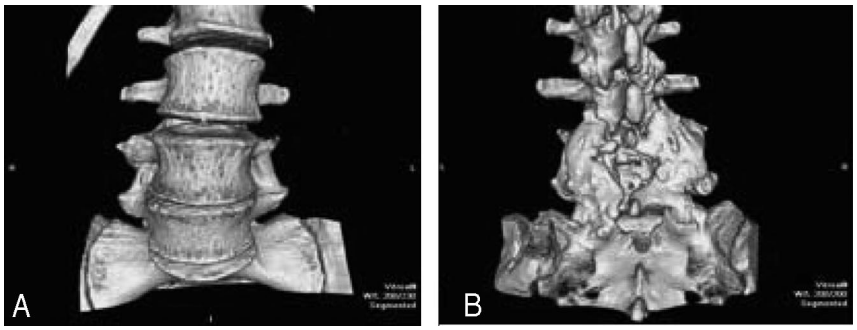


Figure 6. Three-dimensional CT reconstructions in the depiction of bony fusion following posterolateral arthrodesis with OP-1 Putty. AP (**A**) and posteroanterior (**B**) 3D CT scans illustrate the nature of bony fusion following an L45 posterolateral fusion with the use of OP-1 Putty.

nomic benefit to the hospital and/or provider system. Finally, the autograft patients were found to experience mild/moderate donor site pain in 35% of cases at 36+ months, demonstrating that the pain associated with autograft harvest can be both significant and of lasting duration.

There were no significant differences in the occurrence rates of serious adverse events related to device between the study groups at either 24 or 36+ months in the study. There were also no complications or adverse events directly attributable to the OP-1 Putty. Testing for anti-OP-1 antibodies and anti-OP-1 neutralizing antibodies demonstrated transient occurrences between the 6-week and 3-month periods. However, by 24 and 36+ months, no patients had neutralizing antibodies present. It seems based on the data that the formation of anti-OP-1 antibodies does not have any clinically significant effects on either safety or efficacy. Most importantly, the presence of neutralizing antibodies was not correlated with any safety concerns or clinical outcomes.

In terms of the key variable, revision surgery to promote fusion, the overall rates were low in both treatment groups considering the challenging surgical model and the OP-1 group and autograft group showed no significant differences over the course of the study period. At the 36-month follow-up point, 21 of 257 (8.2%) of the OP-1 Putty patients and 11 of 87 (13%) of the autograft patients had undergone further surgery to promote fusion at the index level. At the time of latest follow-up at a mean of over 4 years after the original surgery, only an additional 3 of 144 (2.2%) OP-1 Putty patients and 3 of 58 (5.2%) autograft patients had undergone further surgery for retreatment failure.

■ Conclusion

OP-1 Putty has been designed as an alternative to autograft harvest in posterolateral spinal fusion. In multiple preclinical and early clinical models, OP-1 has produced fusion success results equivalent or superior to that of autograft. Based on this large prospective randomized controlled trial of uninstrumented posterolateral arthrodesis performed for degenerative spondylolisthesis and spinal stenosis, OP-1 Putty is a safe and effective alternative to autograft that results in equivalent overall success outcomes, shorter operative times, and lower intraoperative blood loss, while avoiding the morbidity associated with autograft harvest.

■ Key Points

- OP-1 Putty is a safe and effective alternative to autograft in uninstrumented posterolateral fusion performed for degenerative spondylolisthesis and spinal stenosis.
- At 36+ months, OP-1 Putty resulted in equivalent outcomes in terms of overall success and all clinical and radiographic endpoints.
- The OP-1 Putty group had significantly lower blood loss during surgery and significantly shorter operative times.
- Although antibodies to OP-1 did develop, they resolved in all patients without clinical sequelae.

References

1. Fischgrund JS, Mackay M, Herkowitz HN, et al. 1997 Volvo Award winner in clinical studies. Degenerative lumbar spondylolisthesis with spinal stenosis: a prospective, randomized study comparing decompressive laminectomy and arthrodesis with and without spinal instrumentation. *Spine* 1997;22:2807–12.
2. Herkowitz HN, Kurz LT. Degenerative lumbar spondylolisthesis with spinal stenosis. A prospective study comparing decompression with decompression and intertransverse process arthrodesis. *J Bone Joint Surg Am* 1991;73:802–08.
3. Beirne JC, Barry HJ, Brady FA, et al. Donor site morbidity of the anterior iliac crest following cancellous bone harvest. *Int J Oral Maxillofac Surg* 1996;25:268–71.
4. Cotler JM, Star AM. "Complications of spinal fusion." In: Cotler JM, Cotler HB, eds. *Spine Fusions: Science and Technique*. New York, NY: Springer-Verlag; 1990.
5. Younger EM, Chapman MW. Morbidity at bone graft donor sites. *J Orthop Trauma* 1989;3:192–5.
6. Arrington ED, Smith WJ, Chambers HG, et al. Complications of iliac crest bone graft harvesting. *Clin Orthop Relat Res* 1996;329:300–09.
7. Boden SD. Biology of lumbar spine fusion and use of bone graft substitutes: present, future, and next generation. *Tissue Eng* 2000;6:383–99.
8. Hadjipavlou AG, Simmons JW, Tzermiadianos MN, et al. Plaster of Paris as bone substitute in spinal surgery. *Eur Spine J* 2001;10(suppl 2):189–96.
9. Urist MR. Bone: formation by autoinduction. *Science* 1965;150:893–9.
10. Uludag H, Gao T, Porter TJ, et al. Delivery systems for BMPs: factors contributing to protein retention at an application site. *J Bone Joint Surg Am* 2001;83-A(1 pt 2):128–135.
11. 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.
12. Grauer JN, Patel TC, Erulkar JS, et al. 2000 Young Investigator Research Award winner: evaluation of OP-1 as a graft substitute for intertransverse process lumbar fusion. *Spine* 2001;26:127–33.

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

13. 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.
14. Liao SS, Guan K, Cui FZ, et al. Lumbar spinal fusion with a mineralized collagen matrix and rhBMP-2 in a rabbit model. *Spine* 2003;28:1954–60.
15. Sidhu KS, Prochnow TD, Schmitt P, et al. Anterior cervical interbody fusion with rhBMP-2 and tantalum in a goat model. *Spine J* 2001;1:331–40.
16. Takahashi J, Saito N, Ebara S, et al. Anterior thoracic spinal fusion in dogs by injection of recombinant human bone morphogenetic protein-2 and a synthetic polymer. *J Spinal Disord Tech* 2003;16:137–43.
17. Cook SD, Baffes GC, Wolfe MW, et al. The effect of recombinant human osteogenic protein-1 on healing of large segmental bone defects. *J Bone Joint Surg Am* 1994;76A:827–38.
18. Cook SD, Baffes GC, Wolfe MW, et al., Recombinant human bone morphogenic protein-7 induces healing in canine long bone segmental defect model. *Clin Orthop* 1994;301:302–12.
19. Cook SD, Wolfe MW, Salkeld SL, et al. Effect of recombinant human osteogenic protein-1 on healing of segmental defects in non-human primates. *J Bone Joint Surg Am* 1995;77:734–50.
20. Cook SD, Rueger DC. Osteogenic protein-1. *Clin Orthop* 1996;324:29–38.
21. 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.
22. Paramore CG, Lauryssen C, Rauzzino MJ, et al. The safety of OP-1 for lumbar fusion with decompression: a canine study. *Neurosurgery* 1999;44:1151–5.
23. Magin MN, Delling G. Improved lumbar vertebral interbody fusion using rhOP-1: a comparison of autogenous bone graft, bovine hydroxylapatite (Bio-Oss), and BMP-7 (rhOP-1) in sheep. *Spine* 2001;26:469–78.
24. Furlan JC, Perrin RG, Govender PV, Petrenko Y, et al. Use of osteogenic protein-1 in patients at high risk for spinal pseudarthrosis: a prospective cohort study assessing safety, health-related quality of life, and radiographic fusion. Invited submission from the joint section on disorders of the spine and peripheral nerves, March 2007. *J Neurosurg Spine* 2007;7:486–95.
25. Govender PV, Rampersaud YR, Rickards L, et al. Use of osteogenic protein-1 in spinal fusion: literature review and preliminary results in a prospective series of high-risk cases. *Neurosurg Focus* 2002;13:e4.
26. Vaccaro AR, Patel T, Fischgrund J, et al. A pilot study evaluating the safety and efficacy of OP-1 Putty (rhBMP-7) as a replacement for iliac crest autograft in posterolateral lumbar arthrodesis for degenerative spondylolisthesis. *Spine* 2004;29:1885–92.
27. Vaccaro AR, Anderson DG, Patel T, et al. Comparison of OP-1 Putty (rhBMP-7) to iliac crest autograft for posterolateral lumbar arthrodesis: a minimum 2-year follow-up pilot study. *Spine* 2005;30:2709–16.
28. Vaccaro AR, Whang PG, Patel T, et al. The safety and efficacy of OP-1 (rhBMP-7) as a replacement for iliac crest autograft for posterolateral lumbar arthrodesis: minimum 4-year follow-up of a pilot study. *Spine J* 2008;8:457–65.
29. Mire-Sluis AR, Barrett YC, Devanarayan V, et al., Recommendations for the design and optimization of immunoassays used in the detection of host antibodies against biotechnology product. *J Immunol Methods* 2004;289:1–16.
30. Meyerding HW. Spondylolisthesis. *Surg Gynecol Obstet* 1932;54:371–7.
31. Mooney V. A randomized double-blind prospective study of the efficacy of pulsed electromagnetic fields for lumber interbody fusion. *Spine* 1990;15:708–12.
32. Martin GJ, Boden SD, Morone 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.

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

J Neurosurg: Spine 2:653–657, 2005

Guidelines for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 4: radiographic assessment of fusion

Daniel K. Resnick, M.D., Tanvir F. Choudhri, M.D., Andrew T. Dailey, M.D., Michael W. Groff, M.D., Larry Khoo, M.D., Paul G. Matz, M.D., Praveen Mummaneni, M.D., William C. Watters III, M.D., Jeffrey Wang, M.D., Beverly C. Walters, M.D., M.P.H., and Mark N. Hadley, M.D.

Department of Neurosurgery, University of Wisconsin, Madison, Wisconsin; Department of Neurosurgery, Mount Sinai Medical School, New York, New York; Department of Neurosurgery, University of Washington, Seattle, Washington; Department of Neurosurgery, Indiana University, Indianapolis, Indiana; Departments of Orthopedic Surgery and Neurosurgery, University of California at Los Angeles, California; Department of Neurosurgery, University of Alabama at Birmingham, Alabama; Department of Neurosurgery, Emory University, Atlanta, Georgia; Bone and Joint Clinic of Houston, Texas; and Department of Neurosurgery, Brown University, Providence, Rhode Island

Key Words • **lumbar spine** • **fusion** • **radiography** • **treatment outcome** • **practice guidelines**

Recommendations

Standards. Static lumbar radiographs are not recommended as a stand-alone means to assess fusion status following lumbar arthrodesis surgery.

Guidelines. 1) Lateral flexion and extension radiography is recommended as an adjunct to determine the presence of lumbar fusion postoperatively. The lack of motion between vertebrae, in the absence of rigid instrumentation, is highly suggestive of successful fusion. 2) Technetium-99 bone scanning is not recommended as a means to assess lumbar fusion.

Options. Several radiographic techniques, including static radiography, lateral flexion–extension radiography, and/or CT scanning, often in combination, are recommended as assessment modality options for the noninvasive evaluation of symptomatic patients in whom failed lumbar fusion is suspected.

Rationale

Lumbar fusion is performed in patients with pain due to lumbar degenerative disease. An outcome measure frequently cited in studies evaluating lumbar fusion techniques is the "radiographic fusion rate;" however, radiographic fusion is not consistently defined throughout the literature. The purpose of this review is to examine the literature regarding the ability of various diagnostic techniques to assess fusion status after lumbar fusion is performed to treat degenerative disease.

Search Criteria

A computerized search of the database of the National Library of Medicine between 1966 and July 2003 was conducted using the search terms "lumbar spine fusion assessment," "lumbar spine pseudoarthrosis," or "lumbar spine fusion outcome." The search was restricted to references in the English language involving humans. This yielded a total of 1076 references. The titles and abstracts of each of these references were reviewed. Only papers concerned with the assessment of fusion status following arthrodesis procedures for degenerative lumbar disease were included. Additional articles were obtained from the bibliographies of the selected articles. Forty-five references were identified that provided either direct or supporting evidence relevant to the radiographic assessment of lumbar fusion status. Reports involving Class III or better medical evidence are listed in Table 1. Supportive data are provided by additional references listed in the bibliography.

Scientific Foundation

Open surgical exploration is the only method that allows direct inspection of fusion integrity. This procedure

Abbreviations used in this paper: CT = computerized tomography; NPV = negative predictive value; PPV = positive predictive value; RSA = roentgen stereophotogrammetric analysis.

is considered the gold standard of lumbar fusion assessment.[6,7] It is, therefore, an appropriate benchmark to use in establishing the accuracy and predictive value of noninvasive radiolographic studies for the assessment of fusion status following attempted lumbar fusion surgery.

Plain Radiographs (static)

Anteroposterior and lateral radiographs can demonstrate a continuous bone mass between adjacent vertebral segments following lumbar fusion. Because of their relatively low cost, widespread availability, and long history as a means of assessing fusion, plain spinal radiography remains a common method of assessment of lumbar fusion;[6] however, the limitations of static plain radiography as a reliable test for determining the presence or absence of a solid fusion have been well documented. Brodsky, et al.,[3] reported a 64% correlation between preoperative plain radiographs and surgical exploration in a retrospective study of 214 lumbar fusion exploration procedures in patients who had undergone prior posterolateral fusion. Plain radiography had an 89% sensitivity and 60% specificity for predicting solid fusion. Radiographs interpreted as demonstrating fusion had a PPV of 76%. Those predicting pseudarthrosis had an NPV of 78%. These data indicate a 0.18 likelihood ratio for a false-positive result (chance of a pseudarthrosis discovered at exploration when radiography indicates fusion), and a 2.25 likelihood ratio for a negative test result (chance of a fusion discovered at exploration when the radiography suggests pseudarthrosis).[3] The medical evidence provided by this review is considered Class II for the use of plain lumbar radiography compared with open surgical exploration to assess fusion because of the authors' selection bias for open exploration.

Similarly, in a retrospective study of 75 patients, Kant and coworkers[11] found a positive correlation between static radiography and surgical exploration of lumbar fusion in 68% of their patients (sensitivity 85%, specificity 62%, PPV 76%, and NPV 54%). The likelihood ratio for a positive result was 0.81, and the likelihood ratio for a negative result was 2.24. Finally, in a study of 49 patients treated with posterolateral and posterior interbody fusion with internal fixation, Blumenthal and Gill[1] compared findings on anteroposterior and lateral radiographs (interpreted by two surgeons and two radiologists) with surgical exploration of the fusion mass at the time of reoperation for hardware removal. They reported a 69% agreement between the radiographic diagnosis and surgical findings. The accuracy among the four physicians interpreting the radiographs ranged from 57 to 77% (false-positive rate 42%, false-negative rate 29%). These authors concluded that plain radiography has limited accuracy and validity for the assessment of lumbar fusion. Furthermore, they noted significant intra- and interobserver variation, indicating a lack of reliability (κ 0.4–0.7). Their study provides Class I medical evidence indicating that static radiography is only accurate in determining fusion status in roughly two thirds of cases. Therefore, static anteroposterior and lateral radiographs are not recommended as a stand-alone assessment of the presence of an arthrodesis after lumbar fusion surgery for degenerative disease.

Flexion–Extension Radiography

In 1948 Cleveland, et al.,[6] advocated the use of dynamic lumbar spinal radiography rather than static radiography, for the diagnosis of pseudarthrosis following attempted lumbar fusion surgery. Other authors have also suggested that lateral lumbar flexion–extension radiography allows for appropriate assessment of fusion status.[4] There has been disagreement, however, on the number of allowable degrees of motion at the treated (fused) levels for determining the presence or absence of successful bone fusion.[16]

Brodsky, et al.,[3] compared the findings of lumbar flexion–extension radiography to surgical exploration in a series of 175 patients who underwent reoperation for various indications following instrumented and noninstrumented lumbar fusion. They found a 62% correlation between preoperative flexion–extension radiography and intraoperative findings at exploration (specificity 37%, sensitivity 96%, PPV 70%, and NPV 86%). Their study provides Class II medical evidence that the absence of motion on flexion–extension x-ray films is highly suggestive of a solid fusion. The occurrence of some degree of motion at the treated levels, however, does not necessarily indicate a pseudarthrosis.

Computerized Tomography Scanning

Since the introduction of CT scanning in the 1970s, this modality has been used to assess lumbar fusion. Early studies involved axial sequences alone. Brodsky, et al.,[3] used 6-mm axial slice CT scans and demonstrated a 57% correlation between fusion assessment based on these scans compared with direct surgical exploration in a series of 214 operations on 175 patients. Computerized tomography scanning had a sensitivity of 63%, specificity of 86%, PPV of 72%, and an NPV of 81%. Laasonen and Soini[12] conducted a retrospective review of 20 patients who underwent CT scanning prior to surgical exploration and found an approximate 80% correlation between the CT study–based diagnosis of fusion and intraoperative diagnosis of fusion. Since the publication of these earlier studies, CT imaging technology has advanced. The use of thin-section axial sequences, improved resolution, and multiplanar imaging capability has enhanced the ability of CT scanning to assess lumbar fusion status. There have been no studies comparing these more advanced CT scanning capabilities with direct surgical exploration. Lang and colleagues[14] found that the addition of thin-slice and multiplanar CT scanning resulted in a higher rate of detection of pseudarthrosis compared with plain radiography. Similarly, Chafetz, et al.,[5] demonstrated that direct coronal CT scanning may be more sensitive than two-dimensional reconstructed coronal CT images for the detection of pseudarthrosis. Zinreich and colleagues[21] reported that three-dimensional CT reconstruction may be more sensitive than two-dimensional CT reconstruction for the detection of pseudarthrosis. Siambanes and Mather[20] demonstrated that multiplanar CT imaging detected pseudarthrosis in patients who had undergone posterior lumbar interbody fusion compared with plain radiography that had suggested a solid fusion. Santos and colleagues[18] examined 32 patients who underwent anterior lumbar interbody fusion with carbon fiber cages. Plain static radiographs were interpreted to demonstrate fusion at 86% of the assessed levels. Flexion–extension lumbar radiography suggested fusion rates ranging from 74 to 96% in this same group of patients, depending on the method used to analyze

TABLE 1

*Summary of studies involving radiographic assessment of fusion**

Authors & Year	Class	Description	Comment
Blumenthal & Gill, 1993	I	Retrospective study of 49 patients instrumented lumbar fusion underwent exploration to remove instrumentation. AP & lat radiographs compared w/ op findings w/ 69% agreement. Accuracy ranged among the observers from 57–77%. False–positive rate 42%; false–negative rate 29%.	Limited accuracy of plain x-ray in assessing fusion status w/ low validity (large intra- & interobserver variation).
Bohnsack, et al., 1999	II	Retrospective study of 42 patients (40 lumbar) on utility of planar bone scintigraphy (^{99m}Tc) to assess fusion just before admission for hardware removal. Based on scintigraphy data, pseudarthrosis was suspected in 5 (12%), & the condition was confirmed in 4 during op (10%), 2 diagnosed & 2 undiagnosed. The accuracy of the method was 88%; sensitivity, 50%; specificity, 93%; PPV, 40%; and NPV, 95%. The sensitivity & PPV of bone scintigraphy are low for possible instability after spinal fusion. The method is not sufficient to diagnose pseudarthrosis reliably after arthrodesis.	Based on low sensitivity, bone scan not adequate to diagnose nonunion.
Brodsky, et al., 1991	II	Retrospective study of 214 explorations to remove of internal fixation devices, batteries, or for failed-back surgery in 175 patients w/ PLF. Plain x-rays, polytomography, bending films, &/or CT scans correlated w/ surgical findings. Significant inaccuracy found for all modalities: plain x-rays 36%, polytomograms 41%, bending films 38%, axial CT 43% noncorrelations. Axial CT had lowest inaccuracy (22%), whereas bending films had the highest (27%).	Significant inaccuracy of plain x-ray, polytomograph, bending films, & axial CT in assessing fusion status.
Kant, et al., 1995	II	Retrospective study of 75 patients w/ instrumented lumbar fusions. Single-blinded examiner reviewed x-rays immediately before hardware removal & fusion exploration: 68% correlation btwn radiographic evaluation & intraop observation. Sensitivity 85%, specificity 62%, PPV 76%, & NPV 54%.	Limited accuracy of plain x-rays.
Laasonen & Soini, 1989	II	Retrospective study of 48 patients w/ persistent pain after lumbar fusion examined using CT (6-mm slices, selective sagittal recon). 157 findings observed including: fragmentation of the fusion mass (16), hair-line pseudarthrosis (9), & spinal stenoses (8). Reop in 20 patients: 21 of 27 main lesions detected by CT were confirmed; 6 CT findings were partially or totally incorrect. 16 (80%) of 20 correlations of CT & fusion assessments. 2 cases where CT suggested nonunion but fusion solid at op. 2 cases where CT suggested union w/ pseudarthrosis at op.	Moderate (80%) accuracy of CT in assessing fusion.
Larsen, et al., 1996	II	Prospective study of 25 patients w/ lumbar fusion. All had hardware removal & fusion inspection. Studies to rule out pseudarthrosis included plain radiography, flexion–extension radiography, CT, & bone scintigraphy. Each study evaluated by blinded radiologist. At exploration, instrumentation removed & fusion inspected. No statistically significant correlation was found between radiographic & op findings.	Single-observer blinded study demonstrating no significant correlation btwn radiography & exploration.
Jacobson, et al., 1997	III	Ultra evaluated in 10 patients after posterolat thoracic or lumbar fusion w/in 1 wk before second-look surgery. 20 sites evaluated for bone graft, solid fusion, clefts, fluid collections, & hardware visibility. Ultra & op findings compared. In 3 patients, standard radiographs were reviewed before ultra; blinded ultra evaluation was performed in the remaining 7. Ultra identified all 10 sites of pseudarthrosis seen intraop correctly. Of 10 sites w/ solid fusion at surgery, ultra depicted 6. At 4 sites (2 patients), fusion was mistaken for or obscured by hardware. Overall, sensitivity 100%, specificity 60%, & accuracy 80%.	Class III despite comparison w/ op because of lack of intraobserver reliability data.

* PLF = posterolateral fusion; recon = reconstruction; ultra = ultrasonography.

the x-ray films. The addition of thin-section helical CT scanning reduced the radiographic fusion rate to 65%. The authors concluded that CT scanning is more sensitive than static or flexion–extension lumbar radiography for the detection of pseudarthrosis. Shah, et al.,[19] reached a similar conclusion in their study of 155 patients who underwent posterior lumbar interbody fusion procedures. They found that CT scanning was more sensitive for the detection of abnormalities than plain radiography. These papers are considered to provide Class III medical evidence on the utility of CT scanning for the diagnosis of pseudarthrosis following attempted lumbar fusion.

Technetium-99m Bone Scan

Technetium-99m bone scanning has also been used to assess the integrity of fusion following lumbar arthrodesis surgery. Bohnsack, et al.,[2] performed a retrospective study of 42 patients who underwent lumbar fusion and internal fixation. They obtained ^{99m}Tc bone scans before reoperation for hardware removal. This modality suggested pseudarthrosis in five patients (12%). Pseudarthrosis was found intraoperatively in four patients (10%). In two of these four patients pseudarthrosis was predicted based on the ^{99m}Tc scanning. The accuracy of ^{99m}Tc bone scanning was 88%, its sensitivity was 50%, its specificity was 93%, its PPV was 40%, and its NPV was 95%. This Class II medical evidence suggests that ^{99m}Tc bone scanning is not sufficiently reliable to diagnose pseudarthrosis following a lumbar arthrodesis procedure.[2]

Roentgen Stereophotogrammetric Analysis

Roentgen stereophotogrammetric analysis is a technique that uses radiopaque 0.8-mm tantalum markers implanted into each vertebral level incorporated in the fusion at the time of surgery. The details of the technique have been described elsewhere.[10] Postoperatively, the patient undergoes computerized radiographic assessment in which two 40° angled roentgen tubes are used. Evaluation is performed with the patient in different positions (for example, supine and upright) to detect movement. The technique assesses the amount of movement between the fused vertebral bodies in multiple planes. The amount of allowable movement that determines fusion compared with nonunion, however, is not well defined. This modality has been evaluated in patients at several centers. In a study of 11 patients treated with lumbar fusion, Johnsson and colleagues[10] compared the results of RSA with those of plain radiography at several postoperative time points. In eight patients in whom plain radiography demonstrated successful fusion, RSA revealed a progressive decrease in intervertebral movement over time with achievement of "rigid fusion" within 3 to 12 months. In a follow-up study, Johnsson, et al.,[9] conducted RSA in 12 lumbar fusion patients at multiple postoperative time points. Again, comparative plain radiographs were used and fusion was considered present in all patients. The authors found that in six patients in whom fusion was considered present negligible movement was observed after 1 month postoperatively, whereas in others in whom fusion eventually occurred gradual reduction in intervertebral movement was demonstrated over time. The fact that negligible movement was noted so soon after surgery, when fusion presumably has not yet occurred, is an interesting observation. Pape and associates[17] undertook RSA in 10 patients following lumbar arthrodesis. Based on RSA criteria, fusion was thought to be present in all patients. This finding was confirmed with open surgical exploration in all cases. Although this report supports the accuracy of RSA, because fusion was present in all patients it is not possible to calculate the sensitivity, specificity, PPV, and NPV of RSA compared with exploration from their data.[17]

Other Techniques

Polytomography has been used to assess lumbar fusion status in the pre–CT scanning era, but it has been rarely used since the widespread introduction of CT scanning in the 1970s. In their retrospective study of 214 lumbar fusion exploration procedures in patients who had undergone posterolateral fusion, Brodsky, et al.,[3] found only a 59% correlation of fusion status between preoperative polytomographs and intraoperative findings (sensitivity 65%, specificity 84%, PPV 79%, and NPV 73%). This single study provides Class II medical evidence that polytomography cannot be reliably used to determine the presence of solid osseous arthrodesis following lumbar fusion procedures for degenerative disease.

The use of magnetic resonance imaging to assess for pseudarthrosis following lumbar fusion has been explored by several authors. Lang, et al.,[13] maintained that magnetic resonance imaging added unique information in cases involving lumbar fusion procedures. To date, the importance of this information remains unclear. A single report of the use of ultrasonography to evaluate fusion status was also reviewed.[8] Although the results of this study are promising, the ultrasonography technique has not been rigorously evaluated.

Summary

The assessment of fusion status with static plain radiography is accurate in approximately two thirds of patients treated with lumbar fusion when the radiographic results are compared with surgical exploration findings. Therefore, static plain radiography is not recommended as a stand-alone modality following lumbar fusion procedures. The addition of lateral flexion–extension radiography may improve accuracy because the lack of motion between fused lumbar segments on lateral views is highly suggestive of a solid fusion. Some degree of motion between segments may be present even when the spine has fused. The amount of motion allowable across fused segments is not clear, and the role of internal fixation in limiting motion has also not been adequately addressed. The addition of multiplanar CT scanning results in the detection of pseudarthrosis in some patients in whom fusion has been deemed successful based on plain radiographic criteria. Therefore, CT scanning may be more accurate in the determination of fusion status than plain radiography; however, a rigorous comparison of modern CT scanning and surgical exploration has not been performed. It appears that RSA is exquisitely sensitive for the detection of motion between vertebral bodies, and the loss of motion between treated vertebral segments does appear to indicate the presence of fusion. The modality, however, is invasive and not widely available. Furthermore, the only

comparison of RSA with surgical exploration provided only Class III medical evidence supporting the accuracy of RSA. It is recommended that multiple modalities be used for the noninvasive evaluation of symptomatic patients with suspected fusion failure because no radiographic gold standard exists.

Key Issues for Further Investigation

Modern CT scanning appears to have superior sensitivity compared with plain radiography for the detection of pseudarthrosis. A prospective study of CT scanning findings prior to surgical exploration for instrumentation removal would provide Class I evidence regarding the accuracy of the former compared with the gold standard of surgical exploration. If preoperative flexion–extension radiography is also used, then the influence of internal fixation on the accuracy of flexion–extension radiography could also be addressed.

References

1. Blumenthal SL, Gill K: Can lumbar spine radiographs accurately determine fusion in postoperative patients? Correlation of routine radiographs with a second surgical look at lumbar fusions. **Spine 18:**1186–1189, 1993
2. Bohnsack M, Gosse F, Ruhmann O, et al: The value of scintigraphy in the diagnosis of pseudarthrosis after spinal fusion surgery. **J Spinal Disord 12:**482–484, 1999
3. Brodsky AE, Kovalsky ES, Khalil MA: Correlation of radiologic assessment of lumbar spine fusions with surgical exploration. **Spine 16:**S261–S265, 1991
4. Burkus JK, Dorchak JD, Sanders DL: Radiographic assessment of interbody fusion using recombinant human bone morphogenetic protein type 2. **Spine 28:**372–377, 2003
5. Chafetz N, Cann CE, Morris JM, et al: Pseudarthrosis following lumbar fusion: detection by direct coronal CT scanning. **Radiology 162:**803–805, 1987
6. Cleveland M, Bosworth D, Thompson F: Pseudoarthrosis in the lumbosacral spine. **J Bone Joint Surg Am 30:**302–312, 1948
7. Hilibrand AS, Dina T: The use of diagnostic imaging to assess spinal arthrodesis. **Orthop Clin North Am 29:**591–601, 1998
8. Jacobson JA, Starok M, Pathria MN, et al: Pseudarthrosis: US evaluation after posterolateral spinal fusion: work in progress. **Radiology 204:**853–858, 1997
9. Johnsson R, Axelsson P, Gunnarsson G, et al: Stability of lumbar fusion with transpedicular fixation determined by roentgen stereophotogammetric analysis. **Spine 24:**687–690, 1999
10. Johnsson R, Selvik G, Stromqvist B, et al: Mobility of the lower lumbar spine after posterolateral fusion determined by roentgen stereophotogrammetric analysis. **Spine 15:**347–350, 1990
11. Kant AP, Daum WJ, Dean SM: Evaluation of lumbar spine fusion. Plain radiographs versus direct surgical exploration and observation. **Spine 20:**2313–2317, 1996
12. Laasonen E, Soini J: Low-back pain after lumbar fusion. Surgical and computed tomographic analysis. **Spine 14:**210–213, 1989
13. Lang P, Chafetz N, Genant HK, et al: Lumbar spinal fusion. Assessment of functional stability with magnetic resonance imaging. **Spine 15:**581–588, 1990
14. Lang P, Genant HK, Chafetz N, et al: Three-dimensional computed tomography and multiplanar reformations in the assessment of pseudarthrosis in posterior lumbar fusion patients. **Spine 13:**69–75, 1988
15. Larsen JM, Rimoldi RL, Capen DA, et al: Assessment of pseudarthrosis in pedicle screw fusion: a prospective study comparing plain radiographs, flexion/extension radiographs, CT scanning, and bone scintigraphy with operative findings. **J Spinal Disord 9:**117–120, 1996
16. Mcafee PC, Boden SD, Brantigan JW, et al: Symposium: A critical discrepancy—a criteria of successful arthrodesis following interbody spinal fusions. **Spine 26:**320–334, 2001 (Erratum in **Spine 26:**1103, 2001)
17. Pape D, Fritsch E, Kelm J, et al: Lumbosacral stability after consolidated anteroposterior fusion after instrumentation removal determined by roentgen stereophotogrammetric analysis and direct surgical exploration. **Spine 27:**269–274, 2002
18. Santos ER, Goss DG, Morcom RK, et al: Radiologic assessment of interbody fusion using carbon fiber cages. **Spine 28:** 997–1001, 2003
19. Shah RR, Mohammed S, Saifuddin A, et al: Comparison of plain radiographs with CT scan to evaluate interbody fusion following the use of titanium interbody cages and transpedicular instrumentation. **Eur Spine J 12:**378–385, 2003
20. Siambanes D, Mather S: Comparison of plain radiographs and CT scans in instrumented posterior lumbar interbody fusion. **Orthopedics 21:**165–167, 1998
21. Zinreich SF, Long DM, Davis R, et al: Three-dimensional CT imaging in postsurgical "failed back" syndrome. **J Comput Assist Tomogr 14:**574–580, 1990

Manuscript received December 7, 2004.
Accepted in final form February 18, 2005.
Address reprint requests to: Daniel K. Resnick, M.D., Department of Neurological Surgery, University of Wisconsin Medical School, K4/834 Clinical Science Center, 600 Highland Avenue, Madison, Wisconsin 53792. email: Resnick@neurosurg.wisc.edu.