SPINE Volume 28, Number 22, pp 2527–2530

Inter- and Intraobserver Reliability of Computed Tomography in Assessment of Thoracic Pedicle Screw Placement

Ganesh Rao, MD,* Darrel S. Brodke, MD,† Matthew Rondina, BS,‡ Kent Bacchus, PhD,† and Andrew T. Dailey, MD*

Study Design. Reliability study of computed tomography imaging in 12 cadaver specimens instrumented with titanium or stainless steel thoracic pedicle screws.

Objective. To evaluate inter- and intraobserver reliability of computed tomography scan in determining the accuracy of thoracic pedicle screw placement and to identify the differences in observers' agreement when viewing stainless steel *versus* titanium screws.

Summary of Background Data. Computed tomography is often used to assess the accuracy of pedicle screw placement. Accuracy of screw placement is important in the thoracic spine where pedicle morphometry increases the difficulty of screw placement (vital structures are at increased risk). The current literature lacks a critical evaluation of computed tomography reliability among observers.

Methods. Twelve adult cadavers were instrumented with thoracic pedicle screws. Nine cadavers were instrumented with titanium screws and three with stainless steel screws. The spines were imaged with computed tomography. Three observers used a grading scale to score the extent of pedicle violation and independently scored the placement of each pedicle screw on three separate occasions. Interobserver and intraobserver agreement were determined by using the kappa statistic.

Results. The mean kappa score for interobserver agreement for all 12 specimens (including titanium and stainless steel screws) was 0.51, which correlates with a moderate degree of agreement. Although the interobserver kappa statistics for titanium ($\kappa = 0.53$) and stainless screws ($\kappa = 0.44$) showed a moderate degree of agreement, the intraobserver reliability was substantial ($\kappa = 0.63$). The mean intraobserver kappa for titanium screws was 0.63 and for stainless steel screws was 0.62.

Conclusions. Our data show that interobserver agreement is moderate and intraobserver agreement is substantial when computed tomography is used to assess placement of thoracic pedicle screws. We conclude that computed tomography is reliable for evaluating thoracic pedicle screw placement throughout the thoracic spine.
[Key words: thoracic pedicle screw, computed tomography, interobserver agreement, intraobserver agreement, reliability, titanium screws, stainless steel screws] **Spine 2003;28:2527–2530**

From the Departments of *Neurosurgery and †Orthopedic Surgery, and the ‡School of Medicine, University of Utah, Salt Lake City, Utah.
Acknowledgment date: December 3, 2002. First revision date: February 26, 2003. Acceptance date: April 7, 2003.
The device(s)/drug(s) is/are FDA-approved or approved by corresponding national agency for this indication.
Corporate/Industry funds were received in support of this work. No benefits in any form have been or will be received from a commercial party related directly or indirectly to the subject of this manuscript.
Address correspondence to Andrew T. Dailey, MD, Department of Neurological Surgery, University of Washington, 700 Ninth Avenue, Box 359766, Seattle, WA 98104; E-mail: atdailey@washington.edu

Thoracic pedicle screw fixation has become more common as surgeons have become increasingly familiar with pedicle anatomy and techniques for pedicle fixation.[1–7] Pedicle screws provide rigid fixation, allowing for better deformity correction and greater resistance to pullout than hook and rod constructs.[5,8,9] Because thoracic pedicle screw fixation is now a common procedure, it is important to evaluate the accuracy of screw placement. Instrumentation of the thoracic spine involves unique risks because of the small diameter of thoracic pedicles and the proximity to vital structures such as the spinal cord and aorta. Neurologic complication due to misplaced pedicle screws in the lumbar spine has been well documented.[10,11] Perforation of the pedicle in the thoracic spine can result in a wide variety of undesirable outcomes, including decreased stability of the construct, cerebrospinal fluid leaks, and vascular and neurologic injuries.[5,12–17]

Computed tomography (CT) scan is often used to determine the accuracy of thoracic pedicle screw placement.[5,8,9,14,18,19] Computed tomography has been shown to be useful for assessing the accuracy of screw placement in the lumbar spine[20] and has been shown to reveal up to 10 times as many screws violating the pedicle cortex in the lumbar spine as plain radiographs.[21] However, inter- and intraobserver reliability of CT for assessment of thoracic pedicle screws has not been evaluated. The objective of this study was to determine inter- and intraobserver reliability of CT scan and to investigate the influence of metallic artifact (titanium *vs.* stainless steel screws) on CT scan interpretation.

Materials and Methods

Twelve thoracic spines were harvested from fresh frozen adult cadavers. Nine of the spines were instrumented with titanium thoracic pedicle screws, and three spines were instrumented with stainless steel screws using either a freehand technique[1,5] or stereotactic fluoroscopy. Preinstrumentation CT scans of the spines were obtained to determine the diameter of the thoracic pedicles. Pedicles with an outer cortical diameter <4.5 mm were not instrumented. Pedicles with a diameter of 4.5 to 5.5 mm were instrumented with 4.5 mm Danek M8 screws, and pedicles with a diameter >5.5 mm were instrumented with 5.5 mm Danek Dynalok screws (Sofamor Danek, Memphis, TN). All thoracic levels (T1–T12) were instrumented, provided the pedicles had a diameter >4.5 mm. After the spines were instrumented, they were imaged using CT (General Electric CT/I, Milwaukee, WI). All spines were imaged with the same proto-

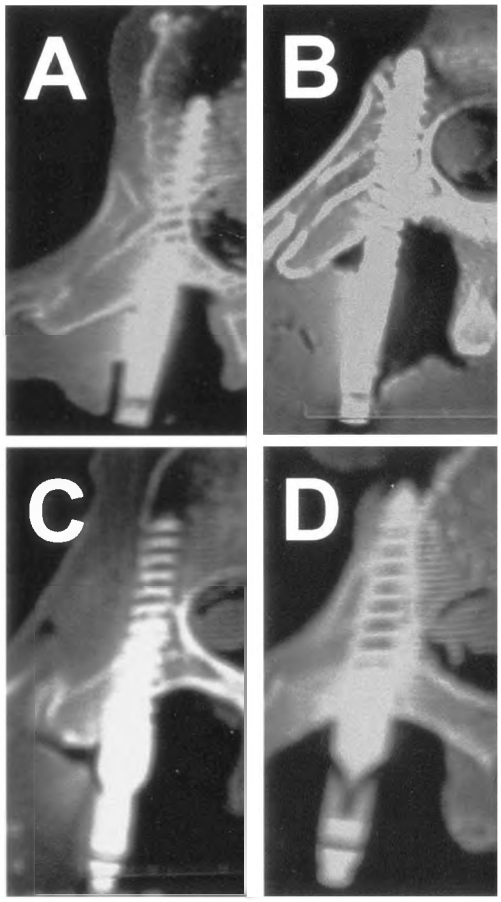

Figure 1. CT scan showing grades for pedicle screw placement. Grade 0 (A), grade 1 (B), grade 2 (C), grade 3 (D).

col of 3 mm axial images with 1 mm overlap. Window settings were adjusted for optimal visualization of the screws and bone.

Three observers (D.S.B., A.T.D., G.R.) then graded each pedicle screw as it appeared on the CT scan on three occasions separated by at least 3 weeks. Thus, each observer made three readings of each set of images for each instrumented spine. Each observer was blinded to the identity of the specimen and the type of pedicle screw used. The grading scale used has been described by our group and has been effectively used for reliability studies of pedicle screw placement.[22] This grading scale ranges from 0 to 3 (Figure 1). Grade 0 is no apparent violation of the pedicle. Grade 1 is <2 mm perforation of the pedicle with one screw thread out of the pedicle. Grade 2 is between 2 and 4 mm of perforation of the pedicle with half of the diameter of the screw outside the pedicle. Grade 3 is >4 mm or complete perforation of the pedicle. No differentiation was made regarding the trajectory of the pedicle violation.

Statistical Analysis. Interobserver and intraobserver kappa scores were then calculated using Stata 7.0 (Stata Corp., College Station, TX). The kappa statistic is a reliability test that measures actual agreement beyond chance compared with potential agreement beyond chance determining[23,24] and has been used in similar studies evaluating interobserver and intraobserver agreement.[25] A kappa of 1 indicates complete agreement of observers in all cases, and a kappa of zero indicates that any observed agreement is attributable only to chance. The most widely used interpretation of the Kappa statistic is the one proposed by Landis and Koch.[23,24,26] In this scheme, the kappa value correlates with a particular degree of agreement (Table 1). Interobserver and intraobserver kappa was calculated for both screw types together and for titanium and stainless steel independently.

Table 1. Relationship Between Kappa Value and Strength of Agreement Between Observers

Value of Kappa	Strength of Agreement
<0	Poor
0–0.20	Slight
0.21–0.40	Fair
0.41–0.60	Moderate
0.61–0.80	Substantial
0.81–1.00	Almost perfect

We subanalyzed interobserver and intraobserver reliability for thoracic levels that have larger pedicles and compared them to thoracic levels that have narrower pedicles. Some investigators have shown that the diameter of the thoracic pedicle narrows between T4 and T8 and is relatively larger between T1 through T3 and T9 through T12.[15,27,28] These narrower pedicles may make interpreting proper pedicle screw placement more difficult as the likelihood of violation of the pedicle is greater. Indeed, Cinotti *et al* suggested that between T4 and T8, pedicle screws of even 4 mm may violate the pedicle.[15] In order to identify whether or not instrumentation of these narrower pedicles would impact interobserver or intraobserver reliability, we compared kappa values between T1–T3 and T9–T12 with T4–T8. Sixty-four pedicles were instrumented from T1 and T3, 92 were instrumented between T4 through T8, and 90 were instrumented between T9 through T12.

■ Results

Two hundred forty-six pedicles were instrumented in total, 182 with titanium screws and 64 with stainless steel screws. Three observers made three readings of each of the 12 instrumented spines. This resulted in 738 total observations per observer. The interobserver kappa value for all screws was 0.51. The interobserver kappa value for titanium screws alone was 0.53. The interobserver kappa value for stainless steel screws was 0.44. These kappa values correlated with a moderate degree of agreement.[24] The mean intraobserver kappa value for all screws was 0.63, which correlates with a substantial degree of agreement. The intraobserver kappa value for titanium screws was 0.63, and the intraobserver kappa value for stainless steel screws was 0.62. The inter- and intraobserver results are summarized in Table 2.

The interobserver kappa value for T1–T3 and T9–T12 combined was 0.51. The interobserver kappa value for T4–T8 was 0.50. These results suggest that the size of

Table 2. Summary of Interobserver Kappa Values

	All Screws	Titanium Screws	Stainless Steel Screws
Interobserver kappa	0.51	0.53	0.44
Intraobserver kappa	0.63	0.63	0.62

the pedicle that was instrumented did not impact the reliability of CT to evaluate screw placement.

■ Discussion

Pedicle screw fixation in the thoracic spine is becoming more common. In the lumbar spine, pedicle screw fixation is facilitated by larger pedicles.[15,27,29–31] The risks of placement of pedicle screws in the thoracic spine have led many surgeons to adapt a position of caution. Perforation rates of the pedicle have been reported at between 1% and 43%,[1] and this variation appears to depend on the postoperative imaging technique (*e.g.*, CT *vs.* plain radiography), the grading scheme used, and the threshold for classifying a screw as misplaced. Although high perforation rates have been reported based on radiographic observation,[14,19] most large series report a <2% complication rate with a low incidence of neurologic and vascular injury and screw revision.[1,9,32]

Computed tomography scan has become the study of choice for postoperative evaluation of pedicle screw placement as well as for evaluating new techniques.[7,20,29,32–36] We recently reported that CT was useful to determine placement of a thoracic pedicle screw when we compared the position of a screw on CT with direct visualization of the instrumented specimen (κ = 0.51, moderate agreement).[22] However, we found that CT had a 62% negative predictive value, indicating that CT may overcall pedicle screw misplacement. Nevertheless, we believe that CT remains the best method to judge screw placement in the clinical environment. We undertook this study to evaluate the reliability of CT scan through the determination of interobserver and intraobserver agreement. In addition, we examined the reliability of CT to assess the accuracy of titanium and stainless steel screws independently.

Our results indicate that for titanium and stainless steel pedicle screws, CT is a reliable study for assessing the accuracy of pedicle screw placement throughout the thoracic spine. The kappa value of 0.51 for all screws placed correlates with a moderate degree of agreement. A moderate degree of agreement has been construed to fall within the fair to good and indicates a level of agreement that is consistent with good reliability.[26,37] In the subanalysis of screw types, the kappa value drops to 0.44 and increases to 0.53 for stainless steel and titanium screws, respectively. The artifact and flare from stainless steel screws may render CT slightly less reliable for determining the accuracy of their placement. This finding is consistent with other studies evaluating the accuracy of CT scan in identifying pedicle screw placement in the lumbar spine. Yoo *et al* found the sensitivity of CT scan in assessing accuracy of pedicle screw placement in the lumbar spine to be 86 ± 5% for titanium screws and only 67 ± 6% for cobalt-chrome screws.[33]

The intraobserver agreement was higher for all screws, titanium and stainless steel (0.63 and 0.62 respectively), and was in the substantial range of agreement. This indicates that each observer performed consistently when grading the appearance of a pedicle screw on CT. This is important because too much variation by each observer would adversely affect the interobserver agreement and decrease the reliability of the test.[25] The substantial degree of intraobserver agreement also validates the grading scheme as a method to evaluate thoracic pedicle screw placement.

In conclusion, CT scan may be used reliably to evaluate the accuracy of pedicle screw placement in the thoracic spine. CT scan will undoubtedly continue to be used to evaluate the accuracy of pedicle screw placement in both clinical and laboratory settings. However, our results show that the type of material used to manufacture a screw used may adversely affect the reliability of CT scan and reduce the ability of observers to accurately assess the placement of the device.

■ Key Points

- There is a need to determine the reliability of CT scan with regard to its use for determining the accuracy of thoracic pedicle screw placement.
- The grading scale presented for extent of pedicle screw perforation has good intraobserver agreement indicating its reliability.
- Computed tomography scan can be used reliably to assess the accuracy of thoracic pedicle screws when titanium screws are used. Stainless screws, however, provide too much artifact for CT scan to be reliable.

Acknowledgments

The authors wish to thank Sofamor Danek for donating the instrumentation used in this study. We would also like to thank John Kestle, MD, for providing guidance with the statistical analysis.

References

1. Belmont PJ Jr, Klemme WR, Dhawan A, et al. *In vivo* accuracy of thoracic pedicle screws. *Spine* 2001;26:2340–6.
2. Chang KW. Oligosegmental correction of post-traumatic thoracolumbar angular kyphosis. *Spine* 1993;18:1909–15.
3. Esses SI, Botsford DJ, Wright T, et al. Operative treatment of spinal fractures with the AO internal fixator. *Spine* 1991;16(suppl):S146–50.
4. Liljenqvist UR, Halm HF, Link TM. Pedicle screw instrumentation of the thoracic spine in idiopathic scoliosis. *Spine* 1997;22:2239–45.
5. Sapkas GS, Papadakis SA, Stathakopoulos DP, et al. Evaluation of pedicle screw position in thoracic and lumbar spine fixation using plain radiographs and computed tomography. A prospective study of 35 patients. *Spine* 1999; 24:1926–9.
6. Suk SI, Lee CK, Kim WJ, et al. Segmental pedicle screw fixation in the treatment of thoracic idiopathic scoliosis. *Spine* 1995;20:1399–405.
7. Wu SS, Hwa SY, Lin LC, et al. Management of rigid post-traumatic kyphosis. *Spine* 1996;21:2260–6; discussion 2267.
8. Liljenqvist U, Hackenberg L, Link T, et al. Pullout strength of pedicle screws versus pedicle and laminar hooks in the thoracic spine. *Acta Orthop Belg* 2001;67:157–63.
9. Suk SI, Kim WJ, Lee SM, et al. Thoracic pedicle screw fixation in spinal deformities: are they really safe? *Spine* 2001;26:2049–57.
10. Blumenthal S, Gill K. Complications of the Wiltse pedicle screw fixation system. *Spine* 1993;18:1867–71.
11. Davne SH, Myers DL. Complications of lumbar spinal fusion with transpedicular instrumentation. *Spine* 1992;17(suppl):S184–9.

12. Ebraheim NA, Jabaly G, Xu R, et al. Anatomic relations of the thoracic pedicle to the adjacent neural structures. *Spine* 1997;22:1553–6; discussion 1557.
13. Kothe R, O'Holleran JD, Liu W, et al. Internal architecture of the thoracic pedicle. An anatomic study. *Spine* 1996;21:264–70.
14. Vaccaro AR, Rizzolo SJ, Allardyce TJ, et al. Placement of pedicle screws in the thoracic spine. Part I: morphometric analysis of the thoracic vertebrae. *J Bone Joint Surg Am* 1995;77:1193–9.
15. Cinotti G, Gumina S, Ripani M, et al. Pedicle instrumentation in the thoracic spine. A morphometric and cadaveric study for placement of screws. *Spine* 1999;24:114–9.
16. Krag MH, Weaver DL, Beynnon BD, et al. Morphometry of the thoracic and lumbar spine related to transpedicular screw placement for surgical spinal fixation. *Spine* 1988;13:27–32.
17. Berry JL, Moran JM, Berg WS, et al. A morphometric study of human lumbar and selected thoracic vertebrae. *Spine* 1987;12:362–7.
18. Steinmann JC, Herkowitz HN, el-Kommos H, et al. Spinal pedicle fixation. Confirmation of an image-based technique for screw placement. *Spine* 1993; 18:1856–61.
19. Vaccaro AR, Rizzolo SJ, Balderston RA, et al. Placement of pedicle screws in the thoracic spine. Part II: an anatomical and radiographic assessment. *J Bone Joint Surg Am* 1995;77:1200–6.
20. Castro WH, Halm H, Jerosch J, et al. Accuracy of pedicle screw placement in lumbar vertebrae. *Spine* 1996;21:1320–4.
21. Farber GL, Place HM, Mazur RA, et al. Accuracy of pedicle screw placement in lumbar fusions by plain radiographs and computed tomography. *Spine* 1995;20:1494–9.
22. Rao G, Brodke DS, Rondina M, et al. Comparison of computerized tomography and direct visualization in thoracic pedicle screw placement. *J Neurosurg* 2002;97(suppl):223–6.
23. Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics* 1977;33:363–74.
24. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74.
25. Christensen FB, Laursen M, Gelineck J, et al. Interobserver and intraobserver agreement of radiograph interpretation with and without pedicle screw implants: the need for a detailed classification system in posterolateral spinal fusion. *Spine* 2001;26:538–43; discussion 543–4.
26. Szklo M, Nieto FJ. *Epidemiology: Beyond the Basics*. Gaithersburg, MD: Aspen; 2000.
27. Ebraheim NA, Xu R, Ahmad M, et al. Projection of the thoracic pedicle and its morphometric analysis. *Spine* 1997;22:233–8.
28. Misenhimer GR, Peek RD, Wiltse LL, et al. Anatomic analysis of pedicle cortical and cancellous diameter as related to screw size. *Spine* 1989;14(4): 367–72.
29. Sjostrom L, Jacobsson O, Karlstrom G, et al. CT analysis of pedicles and screw tracts after implant removal in thoracolumbar fractures. *J Spinal Dis* 1993;6:225–31.
30. Xu R, Ebraheim NA, Ou Y, et al. Anatomic considerations of pedicle screw placement in the thoracic spine. Roy-Camille technique *versus* open-lamina technique. *Spine* 1998;23:1065–8.
31. 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):2279S–2296S.
32. Youkilis AS, Quint DJ, McGillicuddy JE, et al. Stereotactic navigation for placement of pedicle screws in the thoracic spine. *Neurosurgery* 2001;48: 771–8; discussion 778–9.
33. Yoo JU, Ghanayem A, Petersilge C, et al. Accuracy of using computed tomography to identify pedicle screw placement in cadaveric human lumbar spine. *Spine* 1997;22:2668–71.
34. Xu R, Ebraheim NA, Shepherd ME, et al. Thoracic pedicle screw placement guided by computed tomographic measurements. *J Spinal Dis* 1999;12: 222–6.
35. Laine T, Schlenzka D, Makitalo K, et al. Improved accuracy of pedicle screw insertion with computer-assisted surgery. A prospective clinical trial of 30 patients. *Spine* 1997;22:1254–8.
36. Laine T, Makitalo K, Schlenzka D, et al. Accuracy of pedicle screw insertion: a prospective CT study in 30 low back patients. *Eur Spine J* 1997;6:402–5.
37. Fleiss JL. *Statistical Methods for Rates and Proportions*. 2nd ed. New York, NY: Wiley; 1981.