

Observer reliability in evaluating pedicle screw placement using computed tomography

Victor Kosmopoulos · Nicolas Theumann ·
Stefano Binaghi · Constantin Schizas

Received: 14 June 2006 / Accepted: 23 June 2006 / Published online: 12 September 2006
© Springer-Verlag 2006

Abstract Pedicle screw insertion in spinal surgery is a demanding technique with potential risks to neurological structures, for example, within the spinal canal. Assessing screw placement in clinical practice has been performed using plain radiographs and/or mainly axial computed tomography (CT) images. Screw placement using CT image reconstructions in multiple planes has been described, but its reliability has yet to be studied. This study aimed at addressing the clinical issue of interobserver and intraobserver reliability in the use of axial and coronal CT images for the assessment of pedicle screw placement. Fifty nine pedicle screws were studied by two experienced radiologists on two separate occasions. Screw placement was classified as “in”, “out” or “questionable”. On average, 88% and 92% of the screws were classified as “in” by the first and second radiologist, respectively. Intraobserver agreement strength was almost perfect for both observers using either axial or coronal images. Interobserver agreement strength was almost perfect (axial) and substantial (coronal) in the first reading and substantial (axial, coronal) in the second reading. Assessing screw placement in more than one CT imaging plane is not only useful but reliable. Routine use may enhance reporting quality of screw placement by surgeons and radiologists.

Résumé L'insertion des vis pédiculaires dans la chirurgie rachidienne est une technique nécessitant une bonne

expérience, avec des risques potentiels, de lésions des structures neurologiques dans le canal rachidien. Le positionnement des vis pédiculaires peut être évalué à l'aide des radiographies ou du scanner avec reconstruction. Le but de cette étude et de ce travail est de réaliser une étude inter observateurs du positionnement des vis pédiculaires en utilisant le scanner. 59 vis pédiculaires ont été étudiées par des radiologues expérimentés et lors de deux épisodes d'analyse séparés. Le placement des vis a été classé en vis à l'intérieur ou à l'extérieur des pédicules ou en position douteuse. 88 et 92% des vis ont été classées respectivement dans le pédicule par le premier et le deuxième radiologue. Il en est de même des différentes lectures. Evaluer la position des vis pédiculaires à l'aide du CT dans plusieurs plans est non seulement utile mais fiable. Les examens pratiqués de routine pourraient améliorer la qualité des rapports sur le positionnement des vis pédiculaires à la fois pour les chirurgiens et pour les radiologues.

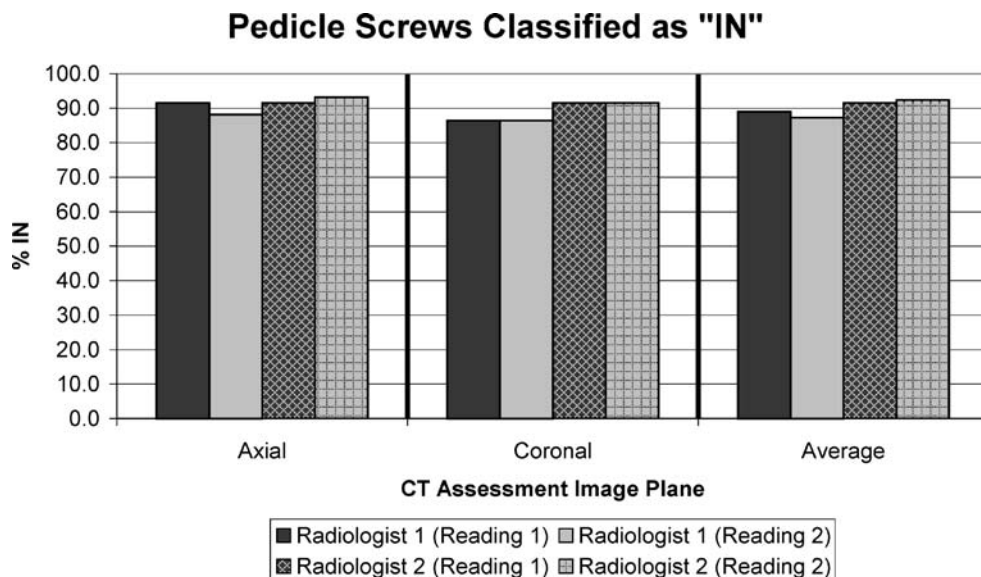
Introduction

Use of pedicle screws for spinal fixation was first reported in the 1970s [25]. Efforts to study optimal placement and evaluate safety of pedicle screw instrumentation in the spine have led to numerous published studies [1–8, 10–16, 18, 19, 21–31]. Specific interests in the accuracy and precision of pedicle screw placement have been driven by safety concerns, including the risk for neurological complications caused by violations to the spinal canal and the possibility for vascular and/or visceral injury [28, 30]. Placement accuracy, however, is based on how it is defined and assessed. A large variation in reported accuracy may exist partly due to the lack of observer reliability in the in vivo post-operative evaluation method and/or the lack of

V. Kosmopoulos (✉) · C. Schizas
Department of Orthopedics,
Hôpital Orthopédique de la Suisse Romande,
Avenue Pierre-Decker 4, CH-1005, Lausanne, Switzerland
e-mail: victor.kosmopoulos@yahoo.com

N. Theumann · S. Binaghi
Department of Radiology,
Centre Hospitalier Universitaire Vaudois,
Lausanne, Switzerland

Fig. 1 Accuracy of pedicle screw placement according to each observer's (radiologist 1 and radiologist 2) reading using axial and coronal computed tomography (CT) images



consensus as to what exactly should be considered as satisfactory pedicle screw placement.

Roentgenogram assessment has been studied and reported to be unreliable as the sole means of evaluating pedicle screw placement [8, 29]. Advances in imaging have led to computed tomography (CT) image assessment methods being regarded as the current gold standard in evaluating in vivo pedicle screw placement success [7, 15]. When compared with radiographs, CT image-based evaluations have been reported to improve accuracy [21]. However, in the majority of in vivo studies, evaluation of pedicle screw placement has been reported using just axial CT images [1–3, 5, 6, 11, 12, 14, 16, 22, 24, 26]. Only a few studies have evaluated multiple CT image planes in reporting placement success [4, 10, 18, 23, 27, 31]. As pedicle screw placement is a three-dimensional (3D) issue, addition of the coronal plane may provide important direction-based information (e.g., cranial or caudal perforations) for the assessment of screw placement.

Interobserver and intraobserver reliability in assessment of in vivo pedicle screw placement using axial and coronal CT images remains to be investigated. This study specifi-

cally aimed to address this issue by having two independent observers evaluate the incidence of misplacement in axial and coronal CT images.

Methods

Fifty-nine titanium screws were implanted in 15 consecutive patients (nine men, six women) having an average age of 41 years during lumbar or lumbosacral fusion. All operations were performed by the same senior surgeon (CS). Post-operative CT scans were obtained for all 15 patients in order to assess implant position. Image acquisition was performed in helical mode using an eight-detector-row Lightspeed CT unit (GE Medical Systems, Milwaukee, WI, USA) in the craniocaudal direction. The imaging series consisted of 2.5-mm-thick CT sections (collimation 8×2.5 mm) at 2-mm intervals with a pitch of 0.875:1 and acquisition parameters of 120 kVp and 280 mAs. The raw data was used to reconstruct axial and coronal cross sections with a field of view adequate for visualisation of the spine. Axial and coronal images were blindly assessed

Table 1 Pedicle screw deviation for screws classified as “out” (OUT) or “questionable” (Q) by each observer during each reading

	Medial	Lateral	Caudal	Cranial	
Reading 1					
Radiologist 1	1 OUT, 2 Q	1 OUT, 1 Q	0	0	Axial
	1 OUT, 3 Q	1 OUT, 3 Q	0	0	Coronal
Radiologist 2	1 OUT, 2 Q	1 OUT, 1 Q	0	0	Axial
	1 OUT	1 OUT, 3 Q	0	0	Coronal
Reading 2					
Radiologist 1	1 OUT, 2 Q	1 OUT, 3 Q	0	0	Axial
	1 OUT, 2 Q	1 OUT, 3 Q	0	1 Q	Coronal
Radiologist 2	1 OUT	1 OUT, 2 Q	0	0	Axial
	1 OUT	1 OUT, 3 Q	0	0	Coronal

Table 2 Interobserver agreement for the first reading

		Radiologist 1				Total
		In	Questionable	Out		
Axial observations						
Radiologist 2	In	53	1	0	54	
	Questionable	1	2	0	3	
	Out	0	0	2	2	
	Total	54	3	2	59	
Coronal observations						
		Radiologist 1				
		In	Questionable	Out	Total	
Radiologist 2	In	49	5	0	54	
	Questionable	2	1	0	3	
	Out	0	0	2	2	
	Total	51	6	2	59	

Axial observations resulted in almost perfect agreement ($\kappa=0.90$) while coronal observations resulted in substantial agreement ($\kappa=0.69$)

separately by two independent radiologists (NT, SB) according to the “in”, “out” or “questionable” classification previously published in the literature [6, 19]. In addition, the observers noted the direction of possible perforation in screws classified as either “questionable” or “out”. The position of these screws on axial images was categorised as either “medial” or “lateral”. On coronal images, “caudal” and “cranial” directions were added to the aforementioned axial categories. Furthermore, each observer evaluated the CT images on two separate occasions (30 days apart) to allow for calculation of intraobserver reliability. Separate weighted kappas (κ) were calculated for each reading of the axial and coronal images to account for the relative importance in disagreement between the three screw placement classifications [17]. Using this approach, a lesser penalty in the final kappa statistic is associated with observer disagreement in classifying a screw as “questionable” than in classifying a screw as “out”. Classification weights were assigned using a quadratic approach [9]. Guidelines for agreement strength were designated by the kappa value. Kappa ranges showing the strength of agreement beyond chance were defined as moderate ($0.41 \leq \kappa \leq 0.60$), substantial ($0.61 \leq \kappa \leq 0.80$) and almost perfect ($0.81 \leq \kappa \leq 1.00$) [17, 20].

Table 3 Interobserver agreement for the second reading

		Radiologist 1				Total
		In	Questionable	Out		
Axial observations						
Radiologist 2	In	51	4	0	55	
	Questionable	1	1	0	2	
	Out	0	0	2	2	
	Total	52	5	2	59	
Coronal observations						
		Radiologist 1				
		In	Questionable	Out	Total	
Radiologist 2	In	50	4	0	54	
	Questionable	1	2	0	3	
	Out	0	0	2	2	
	Total	51	6	2	59	

Axial ($\kappa=0.76$) and coronal ($\kappa=0.78$) observations resulted in substantial agreement.

Results

Averaging the two readings in both the axial and coronal images resulted in 88% and 92% of the pedicle screws being classified as “in” according to the first and second radiologists, respectively. The percentage of pedicle screws classified as “in” for each reading separated by observer can be found in Fig. 1. The observers were consistent in classifying two pedicle screws as “out” (one medial, one lateral) in both the axial and coronal images during both readings (Table 1). During the second reading, one observer found the coronal images to provide additional directional information, classifying one screw as “questionable”, with a possible perforation in the cranial direction.

Interobserver agreement strength (comparing readings between observers) was almost perfect ($\kappa=0.90$) and substantial ($\kappa=0.69$) for screw placement in the first reading of the axial and coronal images, respectively (Table 2). The second reading resulted in a substantial interobserver agreement for both the axial ($\kappa=0.76$) and coronal images ($\kappa=0.78$) (Table 3). Intraobserver agreement strength (comparing chronological readings of a single observer) was almost perfect for both radiologists using either axial or coronal images (Tables 4 and 5).

Table 4 Intraobserver agreement for Radiologist 1

Axial observations		Reading 1			
		In	Questionable	Out	Total
Reading 2	In	52	0	0	52
	Questionable	2	3	0	5
	Out	0	0	2	2
	Total	54	3	2	59
Coronal observations		Reading 1			
		In	Questionable	Out	Total
Reading 2	In	49	2	0	51
	Questionable	2	4	0	6
	Out	0	0	2	2
	Total	51	6	2	59

Axial ($\kappa=0.91$) and coronal ($\kappa=0.84$) observations resulted in almost perfect agreement.

Thirteen of the 15 patients (87%) reported in this study had no screws classified as “out” by either observer regardless of reading. In eight patients (53%) it was agreed by both observers that all pedicle screws were classified as “in” during the first reading. During the second reading, however, one observer found nine (60%) while the other found eleven (73%) patients to have all pedicle screws classified as “in”.

Discussion

Post-operative assessment of in vivo pedicle screw placement is important in order to address safety and efficacy concerns and is helpful in gauging performance and outcome. Advances in imaging technology have made CT the predominant post-operative assessment tool. Most in vivo studies in the literature, however, use only two-dimensional (2D) axial CT images to evaluate the 3D nature of pedicle screw placement [1–3, 5, 6, 11, 12, 14, 16, 22, 24, 26]. The addition of another assessment plane, such as the coronal plane, provides additional information in evaluating placement success. This study aimed at address-

ing the clinical issue of interobserver and intraobserver reliability in the use of axial and coronal CT images for the assessment of pedicle screw placement.

Farber and colleagues using the “in”, “out” or “questionable” classification demonstrated a significant difference between CT and plain radiographic assessment in 16 patients [6]. In their study, they showed no significant difference in interobserver CT readings. Intraobserver comparisons between CT and radiographs, however, did approach the level for statistical significance for all observers. Their findings suggest that CT is more likely to identify marginal or “questionable” screws when compared with plain radiographs. The reliability between CT analysis planes was not discussed in their study.

In a study presented by Fayyazi and colleagues using a simple “in” or “out” classification for pedicle screw placement, the authors found a fair to moderate degree of interobserver and intraobserver agreement using CT and the kappa statistic [7]. The authors of that study, using cadaveric specimens to compare the accuracy of CT with open dissection, concluded that assessment beyond axial CT may be necessary. The use of multiple classifications for the assessment of screw placement, such as the three in

Table 5 Intraobserver agreement for Radiologist 2

Axial observations		Reading 1			
		In	Questionable	Out	Total
Reading 2	In	53	2	0	55
	Questionable	1	1	0	2
	Out	1	1	0	2
	Total	0	0	2	2
Coronal observations		Reading 1			
		In	Questionable	Out	Total
Reading 2	In	53	1	0	54
	Questionable	1	2	0	3
	Out	0	0	2	2
	Total	54	3	2	59

Axial ($\kappa=0.85$) and coronal ($\kappa=0.90$) observations resulted in almost perfect agreement.

our case, and the incorporation of a weighted kappa better represent a clinical setting in which some screws may be very difficult to classify as either “in” or “out” in the view of the observer. This may explain the higher degree of observer agreement found in our study when compared with the aforementioned results presented by Fayyazi and colleagues. However, as in this study, the reliability of results from our study are limited to the classification method used during assessment. This classification method is, however, practical and easy to use in a clinical setting, allowing for the additional flexibility in classifying marginal or “questionable” screws. In addition, the two experienced radiologists in our study were first briefed on the classification to be used prior to assessing pedicle screw placement. Such training of the reader has been shown to be important in improving interobserver variance and CT reading accuracy [13].

Pedicle screw placement is often limited by the diameter of the pedicle. Coronal images allow for full visualisation of this limiting dimension, making them a valuable additional assessment tool along with the currently widely used axial CT images. Our study, to our knowledge, is the first to evaluate observer reliability separately using axial and coronal CT images. Our findings suggest that reconstructed images from these two planes, axial or coronal, are useful and reliable in evaluating pedicle screw placement success. Nevertheless, misinterpretation remains possible. Observer training should be performed to maintain substantial agreement, and the use of both planes is advocated for better assessment of pedicle screw placement success.

References

- Abumi K, Kaneda K, Shono Y, Fujiya M (1999) One-stage posterior decompression and reconstruction of the cervical spine by using pedicle screw fixation systems. *J Neurosurg* 90 (1 Suppl):19–26
- Boachie-Adjei O, Girardi FP, Bansal M, Rawlins BA (2000) Safety and efficacy of pedicle screw placement for adult spinal deformity with a pedicle-probing conventional anatomic technique. *J Spinal Disord* 13(6):496–500
- Carbone JJ, Tortolani PJ, Quartararo LG (2003) Fluoroscopically assisted pedicle screw fixation for thoracic and thoracolumbar injuries: technique and short-term complications. *Spine* 28(1):91–97
- Carl AL, Khanuja HS, Gatto CA, Matsumoto M, vomLehn J, Schenck J, Rohling K, Lorensen W, Vosburgh K (2000) In vivo pedicle screw placement: image-guided virtual vision. *J Spinal Disord* 13(3):225–229
- Castro WH, Halm H, Jerosch J, Malms J, Steinbeck J, Blasius S (1996) Accuracy of pedicle screw placement in lumbar vertebrae. *Spine* 21(11):1320–1324
- Farber GL, Place HM, Mazur RA, Jones DE, Damiano TR (1995) Accuracy of pedicle screw placement in lumbar fusions by plain radiographs and computed tomography. *Spine* 1494–1499
- Fayyazi AH, Hugate RR, Pennypacker J, Gelb DE, Ludwig SC (2004) Accuracy of computed tomography in assessing thoracic pedicle screw malposition. *J Spinal Disord Tech* 17(5):367–371
- Ferrick MR, Kowalski JM, Simmons ED Jr (1997) Reliability of roentgenogram evaluation of pedicle screw position. *Spine* 22 (11):1249–1252, discussion 1253
- Fleiss JL (1981) *Statistical methods for rates and proportions*. Wiley, New York, NY, pp 212–236
- Girardi FP, Cammisa FP Jr, Sandhu HS, Alvarez L (1999) The placement of lumbar pedicle screws using computerized stereotactic guidance. *J Bone Joint Surg Br* 81(5):825–829
- Halm H, Niemeier T, Link T, Liljenqvist U (2000) Segmental pedicle screw instrumentation in idiopathic thoracolumbar and lumbar scoliosis. *Eur Spine J* 9(3):191–197
- Heary RF, Bono CM, Black M (2004) Thoracic pedicle screws: postoperative computerized tomography scanning assessment. *J Neurosurg* 100(4 Suppl Spine):325–331
- Kim HS, Heller JG, Hudgins PA, Fountain JA (2003) The accuracy of computed tomography in assessing cervical pedicle screw placement. *Spine* 28(21):2441–2446
- Kim YJ, Lenke LG, Cheh G, Riew KD (2005) Evaluation of pedicle screw placement in the deformed spine using intraoperative plain radiographs: a comparison with computerized tomography. *Spine* 30(18):2084–2088
- Kosay C, Akcali O, Berk RH, Erbil G, Alici E (2001) A new method for detecting pedicular wall perforation during pedicle screw insertion. *Spine* 26(13):1477–1481
- Kotani Y, Abumi K, Ito M, Minami A (2003) Improved accuracy of computer-assisted cervical pedicle screw insertion. *J Neurosurg* 99(Suppl 3):257–263
- Kundel HL, Polansky M (2003) Measurement of observer agreement. *Radiology* 228:303–308
- Kuntz C, Maher PC, Levine NB, Kurokawa R (2004) Prospective evaluation of thoracic pedicle screw placement using fluoroscopic imaging. *J Spinal Disord Tech* 17(3):206–214
- Laine T, Makitalo K, Schlenzka D, Tallroth K, Poussa M, Alho A (1997) Accuracy of pedicle screw insertion: a prospective CT study in 30 low back patients. *Eur Spine J* 6(6):402–405
- Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* 33:159–174
- Learch TJ, Massie JB, Pathria MN, Ahlgren BA, Garfin SR (2004) Assessment of pedicle screw placement utilizing conventional radiography and computed tomography: a proposed systematic approach to improve accuracy of interpretation. *Spine* 29(7):767–773
- Lee TC, Yang LC, Liliang PC, Su TM, Rau CS, Chen HJ (2004) Single versus separate registration for computer-assisted lumbar pedicle screw placement. *Spine* 29(14):1585–1589
- Papadopoulos EC, Girardi FP, Sama A, Sandhu HS, Cammisa FP Jr (2005) Accuracy of single-time, multilevel registration in image-guided spinal surgery. *Spine J* 5(3):263–267, discussion 268, Jun
- Rampersaud YR, Pik JH, Salonen D, Farooq S (2005) Clinical accuracy of fluoroscopic computer-assisted pedicle screw fixation: a CT analysis. *Spine* 30(7):E183–E190
- Roy-Camille R, Saillant G, Berteaux D, Salgado V, Roy-Camille R, Saillant G, Berteaux D, Salgado V (1976) Osteosynthesis of thoraco-lumbar spine fractures with metal plates screwed through the vertebral pedicles. *Reconst Surg Traumatol* 15:2–16
- Sapkas GS, Papadakis SA, Stathakopoulos DP, Papagelopoulos PJ, Badekas AC, Kaiser JH (1999) Evaluation of pedicle screw position in thoracic and lumbar spine fixation using plain radiographs and computed tomography. A prospective study of 35 patients. *Spine* 24(18):1926–1929

27. Schwarzenbach O, Berlemann U, Jost B, Visarius H, Arm E, Langlotz F, Nolte LP, Ozdoba C (1997) Accuracy of computer-assisted pedicle screw placement. An in vivo computed tomography analysis. *Spine* 22(4):452–458
28. Vaccaro AR, Rizzolo SJ, Balderston RA, Allardyce TJ, Garfin SR, Dolinskas C, An HS (1995) Placement of pedicle screws in the thoracic spine. Part II: an anatomical and radiographic assessment. *J Bone Joint Surg Am* 77:1200–1206
29. Weinstein JN, Spratt KF, Spengler D, Brick C, Reid S (1988) Spinal pedicle fixation: reliability and validity of roentgenogram-based assessment and surgical factors on successful screw placement [see comment]. *Spine* 13:1012–1018
30. Whitecloud TS, Skalley TC, Cook SD, Morgan EL (1989) Roentgenographic measurement of pedicle screw penetration. [Review] [34 refs]. *Clin Orthop* 245:57–68
31. Youkilis AS, Quint DJ, McGillicuddy JE, Papadopoulos SM (2001) Stereotactic navigation for placement of pedicle screws in the thoracic spine. *Neurosurgery* 48(4):771–778, discussion 778–779