

# Interrater Reliability of the Postoperative Epidural Fibrosis Classification: A Histopathologic Study in the Rat Model

Surachai Sae-Jung, Kitti Jirattanaphochai, Chat Sumananont,  
Kriangkrai Wittayapairoj, Kamolsak Sukhonthamarn

*Department of Orthopaedics, Faculty of Medicine, Khon Kaen University, Khon Kaen, Thailand*

**Study Design:** Agreement study.

**Purpose:** To validate the interrater reliability of the histopathological classification of the post-laminectomy epidural fibrosis in an animal model.

**Overview of Literature:** Epidural fibrosis is a common cause of failed back surgery syndrome. Many animal experiments have been developed to investigate the prevention of epidural fibrosis. One of the common outcome measurements is the epidural fibrous adherence grading, but the classification has not yet been validated.

**Methods:** Five identical sets of histopathological digital files of L5–L6 laminectomized adult Sprague-Dawley rats, representing various degrees of postoperative epidural fibrous adherence were randomized and evaluated by five independent assessors masked to the study processes. Epidural fibrosis was rated as grade 0 (no fibrosis), grade 1 (thin fibrous band), grade 2 (continuous fibrous adherence for less than two-thirds of the laminectomy area), or grade 3 (large fibrotic tissue for more than two-thirds of the laminectomy area). A statistical analysis was performed.

**Results:** Four hundred slides were independently evaluated by each assessor. The percent agreement and intraclass correlation coefficient (ICC) between each pair of assessors varied from 73.5% to 81.3% and from 0.81 to 0.86, respectively. The overall ICC was 0.83 (95% confidence interval, 0.81–0.86).

**Conclusions:** The postoperative epidural fibrosis classification showed almost perfect agreement among the assessors. This classification can be used in research involving the histopathology of postoperative epidural fibrosis; for example, for the development of preventions of postoperative epidural fibrosis or treatment in an animal model.

**Keywords:** Fibrosis; Laminectomy; Pathology; Postoperative period

## Introduction

Epidural fibrosis is one of the most common causes of failed back surgery syndrome (FBSS), and it is the caus-

ative factor of pain in up to 36% [1] of FBSS patients. Epidural fibrosis usually results from dural sac compression, nerve root tethering, or from interference with the nerve root vascular supply or cerebrospinal fluid flow [2]. Once

Received Oct 19, 2014; Revised Jan 30, 2015; Accepted Feb 11, 2015

Corresponding author: Surachai Sae-Jung

Department of Orthopaedics, Faculty of Medicine, Khon Kaen University,  
Mittraphar Highway, Muang, Khon Kaen 40002, Thailand

Tel: +66-43-348398, Fax: +66-43-348398, E-mail: [sursea@kku.ac.th](mailto:sursea@kku.ac.th)

the fibrosis forms, there is no effective treatment, and fibrous tissue excision surgery can also cause new epidural fibrosis [3]. One of the modalities for epidural fibrosis is the reduction or prevention of fibrosis formation. Many studies have developed formulations or techniques for its prevention. Most are animal model studies that test the efficacy of the drugs or materials against epidural fibrosis and then compare the study group and a control group based on the histopathology of epidural fibrosis. One of the common postoperative epidural fibrosis outcome measurements is the epidural fibrous adherence grading proposed by He et al. [4]. The classification definition proposed by He et al. [4], with additional definitions for the degree of severity, ranges from normal to a severe degree of epidural fibrosis, and is described as follows: grade 0 (normal, no epidural fibrosis), grade 1 (mild degree, thin fibrous band(s) over dura), grade 2 (moderate degree, continuous adherence observed but less than two-thirds of laminectomy defect) and grade 3 (severe fibrosis, scar tissue adherence large, more than two-thirds of laminectomy area, and/or extending to nerve roots). In the current animal model study, the postoperative epidural fibrous adherence classification was validated for its inter-rater reliability using percent agreement, kappa, and the intraclass correlation coefficient (ICC).

## Materials and Methods

The Animal Research Ethics Committee of Khon Kaen University evaluated and approved the study protocol. Twenty female and 20 male, 8-week-old (adult) Sprague-Dawley rats were laminectomized at their L5–L6 spines. At six weeks after the operation, all of the rats were euthanized using a lethal dose of CO<sub>2</sub>. The L5–L6 vertebrae and all surrounding soft tissues were dissected *en bloc*. The specimens were fixed in phosphate buffered 10% formaldehyde for two days, decalcified in De Castro's fluid, dehydrated in alcohol, embedded in paraffin, and serially cut in a transverse fashion from L5–L6. Then, 5- $\mu$ m cross-sections were processed for hematoxylin and eosin staining [5].

The histopathologic specimens with various degrees of epidural fibrosis were collected. Each of the epidural fibrosis histopathologic slides was labeled with a computer-generated randomized number, which were kept in sealed opaque envelopes. The slides with their number were then scanned with ScanScope FL (Aperio, Vista, CA, USA) into

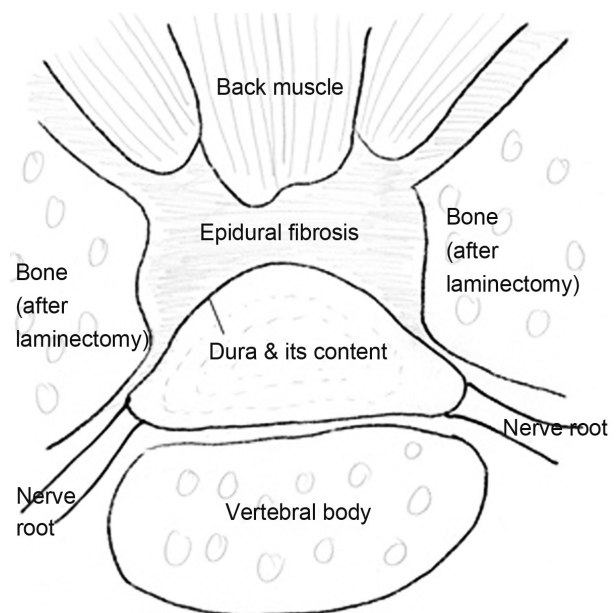
the digital files (ScanScope Virtual Slide). The five identical copies of these files were independently evaluated and graded for the degree of postoperative epidural fibrosis by the five assessors masked to the processes of the study.

### 1. Histopathologic grading of the post-laminectomy epidural fibrosis

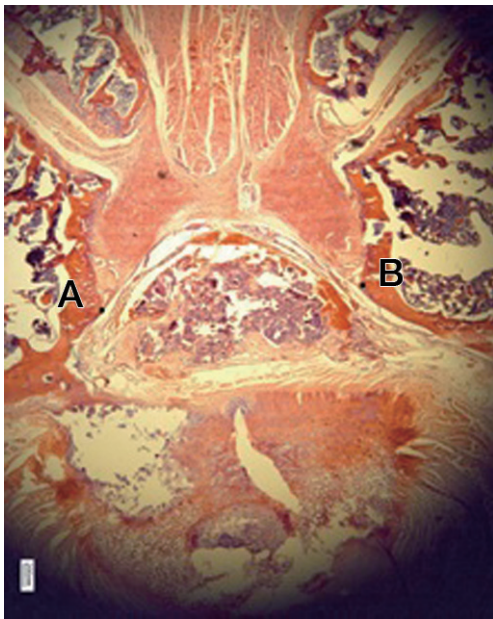
As shown in Fig. 1, the dura and epidural area were assessed. The epidural fibrosis was graded at the area of the posterior aspect of the dura between the left and right rims of the laminectomy site (Fig. 2, point A to B). All assessors were oriented with the epidural fibrosis classification according Table 1 and Figs. 3 and 4. The extent of epidural fibrosis was rated as grade 0 (no fibrosis), grade 1 (mild degree), grade 2 (moderate degree), or grade 3 (severe degree) [4].

### 2. Sample size determination

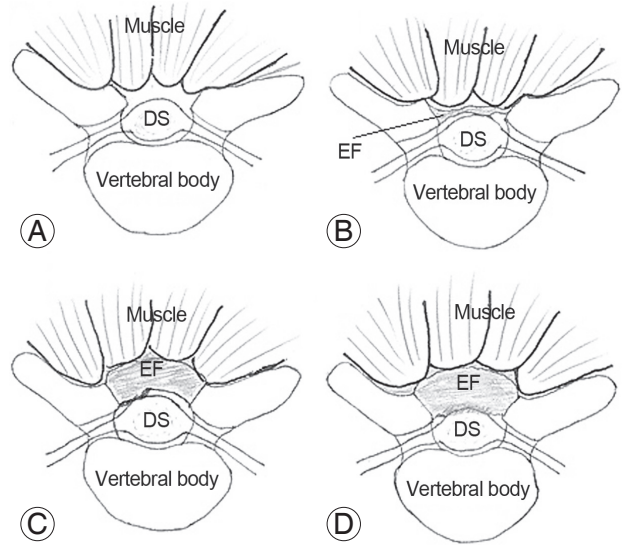
For the 2-rater reliability study, 400 histopathological slides were needed to ensure that the estimated percent agreement fell within 5% of its true error-free value [6]. For a high-powered study, 5 raters were needed to ensure 80% of the predetermined percentage of agreement [6].



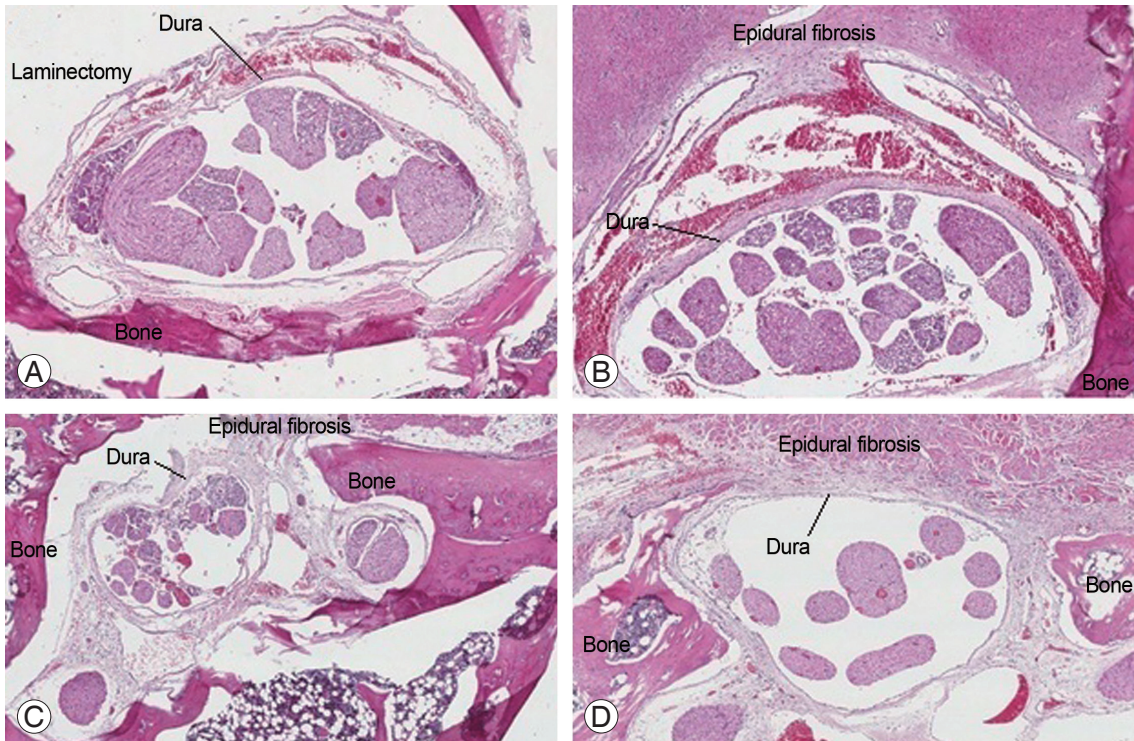
**Fig. 1.** The schematic drawing of the cross-section of the laminectomized lumbar spine showing epidural fibrosis. The epidural fibrosis adheres to the dura in a curvilinear pattern.



**Fig. 2.** Photograph showing post-laminectomy epidural fibrosis (between points A and B of the laminectomy site) (H&E, x2).



**Fig. 3.** Schematic drawing showing post-laminectomy epidural fibrosis grading. **(A)** Grade 0: normal, no fibrosis around the dura. **(B)** Grade 1: mild degree, only thin fibrous band over the dura. **(C)** Grade 2: moderate degree, continuous fibrosis adheres to the dura for less than two-thirds of the laminectomy defect. **(D)** Grade 3: severe degree, large scar tissue adheres to the dura for more than two-thirds of the laminectomy area. DS, dural sac; EF, epidural fibrosis.



**Fig. 4.** Photograph showing histopathology of epidural fibrosis (H&E, x2). **(A)** Grade 0: normal, no fibrosis around the dura. **(B)** Grade 1: mild degree, only thin fibrous band over the dura. **(C)** Grade 2: moderate degree, continuous fibrosis adhere the dura for less than two-thirds of the laminectomy defect. **(D)** Grade 3: severe degree, large scar tissue adhere dura for more than two-thirds of the laminectomy area.

**Table 1.** Histopathological classification of post-laminectomy epidural fibrosis

Grade	Degree	Definition
0	No fibrosis	Dura mater free of fibrotic tissue
1	Mild	Only thin fibrous band(s) over dura observed
2	Moderate	Continuous adherence observed but less than two-thirds of laminectomy defect
3	Severe	Scar tissue adherence large, more than two-thirds of laminectomy area, and/or extending to nerve roots

**Table 2.** Epidural fibrosis grading from all 5 assessors

Assessor	Epidural fibrosis			
	Grade 0	Grade 1	Grade 2	Grade 3
A	7	79	131	183
B	4	89	118	189
C	5	80	128	187
D	8	84	128	180
E	5	87	123	185

### 3. Statistical methods

The obtained results of histopathological grading were collected, and analyzed. The percent agreement, kappa, ICC and 95% confidence interval (CI) were calculated. All data were analyzed with STATA ver. 11.2 (StataCorp, College Station, TX, USA) and IBM SPSS ver. 20 (IBM Co., Armonk, NY, USA).

## Results

Most (77.6%) of the distribution of epidural fibrosis grading by the assessors was between moderate and severe (Table 2). The percent agreement between the two assessors varied from 73.5% to 81.3%. The ICC ranged between 0.81 and 0.86 (Table 3).

The overall ICC of all 5 assessors was 0.83 (95% CI, 0.81–0.86). Grade 2 had a lower kappa (0.51) than the other grades, which were between 0.64 and 0.87 (Table 4).

## Discussion

Due to the difficulty in obtaining human tissue for evaluation of postoperative epidural fibrosis, the animal model has an important role in such investigations. Another advantage of using an animal model is that the disease process (or pathological changes) is accelerated in animals, so

it is possible to study the process in shorter time than in humans [7]. For post-laminectomy epidural fibrosis, the animal model used in our study was the Sprague-Dawley rat, which is relatively small but large enough to provide a suitable structure for a laminectomy and wound healing study.

Anatomical dissection of the rat's spine reveals an epidural space, basement membrane, sweat gland, dermis, and adipose tissue, similar to that found in humans [8]. The rat's vertebral endplate has an increasing diameter from the cervical to the lumbar spine—similar to humans, but unlike large quadrupeds—in which the spine diameter stays much the same over the whole length of the spine [9]. Consequently, we postulate that the epidural fibrosis seen in the rat model can be generalized to humans.

LaRocca and Macnab [10] demonstrated in a laminectomized dog model that peridural fibrosis could be identified in both the epidural space and the nerve roots when both are exposed. Nerve root fibrosis depended on the laminectomy procedure and whether it reached the nerve root. In the current study, only the epidural space was exposed after the laminectomy, so only the epidural area was evaluated for postoperative peridural fibrosis grading. Sen et al. [11] described that the epidural fibrosis tended to form a curvilinear pattern surrounding the dural sac following the contour of the inner laminar surface. In the laminectomy model, therefore, we rated the epidural

**Table 3.** Percentage agreement, intraclass correlation coefficient and 95% confidence interval of intraclass correlation coefficient between the assessors

Assessor	A	B	C	D	E
<b>A</b>					
Agreement (%)	-	78.8	81.3	75.8	81.0
Intraclass correlation coefficient	-	0.84	0.86	0.82	0.86
95% Confidence interval	-	0.82–0.86	0.84–0.88	0.80–0.84	0.84–0.88
<b>B</b>					
Agreement (%)	78.8	-	79.5	78.5	79.0
Intraclass correlation coefficient	0.84	-	0.83	0.83	0.84
95% Confidence interval	0.82–0.86	-	0.82–0.84	0.81–0.85	0.82–0.86
<b>C</b>					
Agreement (%)	81.3	79.5	-	75.8	77.8
Intraclass correlation coefficient	0.86	0.83	-	0.81	0.83
95% Confidence interval	0.84–0.88	0.82–0.84	-	0.80–0.82	0.82–0.86
<b>D</b>					
Agreement (%)	75.8	78.5	75.8	-	73.5
Intraclass correlation coefficient	0.82	0.83	0.81	-	0.81
95% Confidence interval	0.80–0.84	0.81–0.85	0.80–0.82	-	0.78–0.83
<b>E</b>					
Agreement (%)	81.0	79.0	77.8	73.5	-
Intraclass correlation coefficient	0.86	0.84	0.83	0.81	-
95% Confidence interval	0.84–0.88	0.82–0.86	0.82–0.86	0.78–0.83	-

**Table 4.** Kappa of each epidural fibrosis grade from all assessors

Grade	Kappa
0	0.81
1	0.87
2	0.51
3	0.64

fibrosis from the left to the right rim of the laminectomy site (Figs. 1–3).

According to the histopathological classification of postoperative epidural fibrosis in animal models, the classification by He et al. [4] is the most commonly used for prevention research [12–22]. Most studies used a rat model, while some used rabbit and dog models. The classification is based upon the opinion of the pathologist who rates the degree of epidural fibrous adherence. Sucu et al. [23] found no validated histopathological grading system for epidural fibrosis. Therefore, the histopatho-

logical classification for postoperative epidural fibrosis experimental research should be validated and tested for its interrater reliability between the assessors who rate the postoperative epidural fibrosis. The kappa statistics or the chance agreement correction statistics are better than percentage agreement for describing interrater reliability [6]. Concerning the agreement statistics of the ordinal data, the weighted kappa or ICC [24] is preferable to Cohen's kappa [25,26]. We used the quadratic-weighted kappa or ICC [24] because we considered partial agreement to be when the pathologists differently rated the epidural fibrosis for 1 or 2 grades and considered disagreement to be when the assessors differently rated the epidural fibrosis by more than 2 grades (e.g., one assessor rated grade 0 fibrosis while another rated grade 3 fibrosis). In the current study, therefore, we use the ICC for describing the interrater reliability instead of Cohen's kappa.

As mentioned above, the ICC (equivalent to the weighted kappa) [24] was suitable for reporting the interrater reliability of the epidural fibrosis classification, which was

**Table 5.** Interpretation of kappa and intraclass correlation coefficient according to Fleiss [24] and Viera and Garrett [29]

Coefficient value	Interpretation
<0	Less than chance agreement
0.01–0.20	Slight agreement
0.21–0.40	Fair agreement
0.41–0.60	Moderate agreement
0.61–0.80	Substantial agreement
0.81–0.99	Almost perfect agreement

the ordinal parameter [27]. Weighted kappa credits the partial agreement between the raters. However, to test the reliability of each grade of epidural fibrosis, we adhered rigidly to the definition of each epidural fibrosis grade (not accepting partial agreement) and report the kappa (Table 4), instead of the weighted kappa or ICC.

According to the kappa statistics, Landis and Koch [28] suggested that a kappa (a) >0.75 (excellent agreement), (b) <0.40 (poor agreement), and (c) 0.40–0.75 (intermediate to good agreement). Viera and Garrett [29] further classified the interpretation of kappa (Table 5). The overall ICC of the epidural fibrous adherence in this study was 0.83 ( $p < 0.001$ ; 95% CI, 0.81–0.86). These represented the excellent interrater reliability of the classification for post-operative epidural fibrous adherence.

Grade 0 or 1 epidural fibrosis showed almost perfect agreement among the assessors, grade 2 indicated moderate agreement (kappa 0.51) and grade 3 had better agreement (Table 4). Grade 0 (no fibrosis), grade 1 (minimal fibrosis), and grade 3 (severe fibrosis) were not difficult to rate, while grade 2 (moderate fibrosis) was more difficult to rate since the ratings varied from mild to severe. We suggest that grade 2 should be rated with caution, and that consultation may be needed in equivocal cases.

We validated the epidural fibrosis classification based on histopathology in an animal model. For clinical relevance, Ross et al. [30] described the association between the amount of epidural fibrosis and recurrent radicular pain following lumbar discectomy. In their study, all patients underwent magnetic resonance imaging (MRI) at six months postoperatively. The amount of fibrosis was rated as no scar, mild, moderate, and severe scar. They reported that the patients who had severe epidural fibrosis were 3.2 times more likely to experience recurrent radicular symptoms than patients who had lesser degrees of fibrosis. Sen

et al. [11] demonstrated in their clinical and MRI studies that the patients who had lesser degree of epidural fibrosis showed better clinical outcomes. We supposed that this epidural classification can be used in humans. However, because of the ethical issue and the difficulty in obtaining human tissue for the assessment of post-laminectomy epidural fibrosis, the classification definition can be applied to other investigational modalities, such as the axial image of the spinal MRI.

## Conclusions

The histopathological classification of post-laminectomy epidural fibrosis showed excellent interrater reliability. The classification can be used for the assessment of epidural fibrosis and, certainly, for experimental animal model research for epidural fibrosis prevention or treatment.

## Conflict of Interest

No potential conflict of interest relevant to this article was reported.

## Acknowledgments

We thank Dr. Sakda Waraasawapati for his support regarding the histopathological consultation and Mr. Bryan Roderick Hamman for assistance with the English-language presentation of the manuscript.

## References

1. Chan CW, Peng P. Failed back surgery syndrome. *Pain Med* 2011;12:577-606.
2. Jayson MI. The role of vascular damage and fibrosis in the pathogenesis of nerve root damage. *Clin Orthop Relat Res* 1992;(279):40-8.
3. Liu S, Boutrand JP, Bittoun J, Tadie M. A collagen-based sealant to prevent in vivo reformation of epidural scar adhesions in an adult rat laminectomy model. *J Neurosurg* 2002;97:69-74.
4. He Y, Revel M, Loty B. A quantitative model of post-laminectomy scar formation: effects of a nonsteroidal anti-inflammatory drug. *Spine (Phila Pa 1976)* 1995; 20:557-63.
5. Sae-Jung S, Jirattanaphochai K. Prevention of peridural fibrosis using a cyclooxygenase-2 inhibitor

- (nonsteroidal anti-inflammatory drug) soaked in absorbable gelatin sponge: an experimental comparative animal model. *Spine (Phila Pa 1976)* 2013;38:E985-91.
- Gwet KL. *Handbook of inter-rater reliability: the definitive guide to measuring the extent of agreement among multiple raters*. 3rd ed. Gaithersburg, MD: Advanced Analytics, LLC; 2012.
  - Mogford JE, Mustoe TA. Experimental models of wound healing. In: Falanga V, editor. *Cutaneous wound healing*. London: Martin Dunitz Ltd; 2001. p.109-22.
  - Dorsett-Martin WA. Rat models of skin wound healing: a review. *Wound Repair Regen* 2004;12:591-9.
  - Alini M, Eisenstein SM, Ito K, et al. Are animal models useful for studying human disc disorders/degeneration? *Eur Spine J* 2008;17:2-19.
  - LaRocca H, Macnab I. The laminectomy membrane: studies in its evolution, characteristics, effects and prophylaxis in dogs. *J Bone Joint Surg Br* 1974;56:545-50.
  - Sen O, Kizilkilic O, Aydin MV, et al. The role of closed-suction drainage in preventing epidural fibrosis and its correlation with a new grading system of epidural fibrosis on the basis of MRI. *Eur Spine J* 2005;14:409-14.
  - Kasimcan MO, Bakar B, Aktas S, Alhan A, Yilmaz M. Effectiveness of the biophysical barriers on the peridural fibrosis of a postlaminectomy rat model: an experimental research. *Injury* 2011;42:778-81.
  - Yildiz KH, Gezen F, Is M, Cukur S, Dosoglu M. Mitomycin C, 5-fluorouracil, and cyclosporin A prevent epidural fibrosis in an experimental laminectomy model. *Eur Spine J* 2007;16:1525-30.
  - Kurt G, Cemil B, Celik B, Durdag E, Erdem O, Ceviker N. Comparison of Oxiplex and Gore-Tex effectiveness in an experimental peridural fibrosis model. *Neurocirugia (Astur)* 2009;20:360-6.
  - Emmez H, Kardes O, Dogulu F, Kurt G, Memis L, Baykaner MK. Role of antifibrotic cytokine interferon-gamma in the prevention of postlaminectomy peridural fibrosis in rats. *Neurosurgery* 2008;62:1351-7.
  - Shih HN, Fang JF, Chen JH, et al. Reduction in experimental peridural adhesion with the use of a crosslinked hyaluronate/collagen membrane. *J Biomed Mater Res B Appl Biomater* 2004;71:421-8.
  - Dogulu F, Kurt G, Emmez H, et al. Topical mitomycin C-induced inhibition of postlaminectomy peridural fibrosis in rabbits. *J Neurosurg* 2003;99:76-9.
  - Gerszten PC, Moossy JJ, Flickinger JC, Gerszten K, Kalend A, Martinez AJ. Inhibition of peridural fibrosis after laminectomy using low-dose external beam radiation in a dog model. *Neurosurgery* 2000;46:1478-85.
  - Kaya M, Yildirim CH, Kosemehmetoglu K, et al. Alpha-lipoic acid reduces peridural fibrosis after laminectomy of lumbar vertebrae in rabbits. *Acta Neurochir (Wien)* 2012;154:1241-5.
  - Tatsui CE, Martinez G, Li X, Pattany P, Levi AD. Evaluation of DuraGen in preventing peridural fibrosis in rabbits: invited submission from the Joint Section Meeting on Disorders of the Spine and Peripheral Nerves, March 2005. *J Neurosurg Spine* 2006;4:51-9.
  - Bora H, Aykol SV, Akyurek N, Akmansu M, Ataoglu O. Inhibition of epidural scar tissue formation after spinal surgery: external irradiation vs. spinal membrane application. *Int J Radiat Oncol Biol Phys* 2001;51:507-13.
  - Karatay M, Celik H, Koktekir E, et al. Role of tenoxicam in the prevention of postlaminectomy peridural fibrosis in rats. *J Neurol Sci Turk* 2013;30:559-65.
  - Sucu HK, Sevin IE, Rezanko T, Hamdi B, Sevin G. Prevention of anterior scar formation following discectomy with a MediShield adhesion barrier: randomized experimental trial. *Turk Neurosurg* 2013;23:317-22.
  - Fleiss JL. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educ Psychol Meas* 1973;33:613-9.
  - Schuster C. A note on the interpretation of weighted kappa and its relations to other rater agreement statistics for metric scales. *Educ Psychol Meas* 2004;64:243-53.
  - Warren MJ. Cohen's weighted kappa with additive weights. *Adv Data Anal Classif* 2013;7:41-55.
  - Jakobsson U, Westergren A. Statistical methods for assessing agreement for ordinal data. *Scand J Caring Sci* 2005;19:427-31.
  - Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-74.
  - Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Med* 2005;37:

- 360-3.
30. Ross JS, Robertson JT, Frederickson RC, et al. Association between peridural scar and recurrent radicular pain after lumbar discectomy: magnetic resonance evaluation. ADCON-L European Study Group. *Neurosurgery* 1996;38:855-61.