Clinical Article

Interobserver and Intraobserver Reliability of Sub-Axial Injury Classification and Severity Scale between Radiologist, Resident and Spine Surgeon

Woo Jin Lee, M.D.,[1] Seung Hwan Yoon, M.D., Ph.D.,[1] Yeo Ju Kim, M.D., Ph.D.,[2] Ji Yong Kim, M.D.,[3] Hyung Chun Park, M.D., Ph.D.,[1] Chon Oon Park, M.D., Ph.D.[1]

Departments of Neurosurgery,[1] Radiology,[2] Inha University College of Medicine, Incheon, Korea
Department of Neurosurgery,[3] Korea Armed Forces Busan Hospital, Busan, Korea

Objective : The sub-axial injury classification (SLIC) and severity scale was developed to decide whether to operate the cervical injured patient or not, but the reliability of SLIC and severity scale among the different physicians was not well known. Therefore, we evaluated the reliability of SLIC among a spine surgeon, a resident of neurosurgery and a neuro-radiologist.
Methods : In retrograde review in single hospital from 2002 to 2009 years, 75 cases of sub-axial spine injured patients underwent operation. Each case was blindly reviewed for the SLIC and severity scale by 3 different observers by two times with 4 weeks interval with randomly allocated. The compared axis was the injury morphology score, the disco-ligamentous complex score, the neurological status score and total SLIC score; the neurological status score was derived from the review of medical record. The kappa value was used for the statistical analysis.
Results : Interobserver agreement of SLIC and severity scale was substantial agreement in the score of injury morphology [intraclass correlation (ICC)=0.603] and total SLIC and severity sacle (ICC value=0.775), but was fair agreement in the disco-ligamentous complex score (ICC value=0.304). Intraobserver agreements were almost perfect agreement in whole scales with ICC of 0.974 in a spine surgeon, 0.948 in a resident of neurosurgery, and 0.963 in a neuro-radiologist.
Conclusion : The SLIC and severity scale is comprehensive and easily applicable tool in spine injured patient. Moreover, it is very useful tool to communicate among spine surgeons, residents of neurosurgery and neuro-radiologists with sufficient reproducibility.

Key Words : Cervical trauma · Sub-axial injury classification scale · Cervical spine injury · Interobserver agreement · Intraobserver reliability.

INTRODUCTION

Despite technological advances in spine surgery, classification of sub-axial cervical spine injuries remains largely descriptive, lacking standardization and any relationship to prognosis or clinical decision making[10]. The sub-axial injury classification (SLIC) and severity scale was developed by Vaccaro et al.[10] to define a classification system for sub-axial cervical spine trauma that conveys information about injury pattern and severity as well as treatment considerations and prognosis, and it was world-widely used[1,5,6,10]. The reproducibility and reliability of SLIC scale are very important for making decision to operate and communicating among physicians. Unfortunately, there was only a few study results reported about this topic in the literature review[7]. Therefore, we conducted this study to evaluate the reliability of SLIC among spine surgeons, residents of neurosurgery and neuro-radiologists.

MATERIALS AND METHODS

This study was designed as retrospective review of sub-axial spine injured patient with surgical procedure. A total 95 cases of sub-axial spine injury were underwent surgery from April 2002 to December 2009. In the review of the images, the poor quality of radiological findings was observed in 20 cases, and they were dropped out from this study. Finally, 75 cases with

medical record, computed tomography (CT) and magnetic resonance imaging (MRI) were recruited as subjects in this study. The mean age of these subjects was 56 years (18-82 years), and the 56 (74.7%) were males and 19 (25.3%) female.

We classified the cervical injured patient according to the comprehensive classification system for SLIC and severity scale which had been developed by Vaccaro et al.[11] SLIC scale suggests three major categories as indicators, which would direct the treatment of subaxial injuries. These are : 1) injury morphology as determined by the pattern of spinal column disruption on available imaging studies, 2) integrity of the disco-ligamentous complex (DLC) represented by both anterior and posterior ligamentous structures as well as the intervertebral disc, and 3) neurological status of the patient (Table 1). These three categories are widely recognized as predictors of clinical outcome and influenced on decision making of treatment. Within each of the three categories, subgroups have been identified and graded from the least to the most severe (Table 1).

The SLIC and severity scale is a useful classification system for sub-axial cervical trauma, incorporating pertinent characteristics for generating prognoses and courses of management[9]. A numeric value is generated from each axis (injury morphology, integrity of the disco-ligamentous complex, and neurological status), specific to the descriptive identifier. Injury patterns that are known to result in worse outcomes or require surgical intervention (spinal instability, neurological injury) are weighted to receive greater point values. These three numbers, one from each axis, are summed to provide an overall SLIC score and severity scale. The higher the number of points assigned to a particular category, the more severe the injury and the more likely a surgical procedure is indicated. In this study, the surgical procedure was indicated in more than 5 points of SLIC and severity scale, the non-surgical treatment was done in lesser than 3 points of SLIC and severity scale. A score of four is considered equivocal[4]. All cases included in this study were all surgically treated cases with more than 5 points of SLIC and severity scale.

The score of SLIC and severity scale was retrospectively examined by three different observers by carefully reviewing the preoperative simple radiographs, preoperative CT, preoperative MRI and the medical records. The observers were consisted with a spine surgeon, a resident of neurosurgery and a neuro-radiologist. The patient information was not given to the observers and the list of pa-

tient was randomly allocated. The measurement of neurological status was derived from the given medical record. The score of SLIC and severity scale was double examined with 4 weeks interval to check the intraobserver reliability.

Relative to observers as data consistency has been assessed through two approaches : between-observer reliability (interobserver agreement) and within-observer reliability (intraobserver reliability). The interobserver agreement and the intraobserver reliability were checked from each axis and total sum of the SLIC and severity scale, and evaluated with intraclass correlation (ICC) with Cohen's kappa. The interpretation of Cohen's kappa was followed the guideline of Viera and Garrett[12] : $\kappa<0.20$, slight agreement; $0.21<\kappa<0.40$, fair agreement; $0.41<\kappa<0.60$, moderate agreement; $0.61<\kappa<0.80$, substantial agreement; $0.81<\kappa<0.99$, almost perfect agreement. Statistical comparisons were analyzed with SPSS 12.0 (SPSS, Chicago, IL, USA).

RESULTS

Total 450 cases of sub-axial spine injured patient who underwent the surgical procedure were enrolled in this study, by double checking the 75 cases from three different observers. The interobserver agreement and the intraobserver reliability of each axis and total sum of the SLIC and severity scale was summa-

Table 1. Subaxial injury classification scale

	Points
Morphology	
No abnormality	0
Compression	1
Burst	+1=2
Distraction (e.g., facet perch, hyperextension)	3
Rotation/translation (e.g., facet dislocation, unstable teardrop or advanced staged flexion compression injury)	4
Disco-ligamentous complex	
Intact	0
Indeterminate (e.g., isolated interspinous widening, MRI signal change only)	1
Disrupted (e.g., widening of disc space, facet perch or dislocation)	2
Neurological status	
Intact	0
Root injury	1
Complete cord injury	2
Incomplete cord injury	3
Continuous cord compression in setting of neuro deficit (neuro modifier)	+1

MRI : magnetic resonance image

Table 2. The interobserver agreement and intraobserver reliability of sub-axial injury classification (SLIC) and severity scale : mean (conficient interval)

	Interobserver agreement	Intraobserver reliability
Injury morphology	0.603 (0.455-0.731)	0.921 (0.885-0.952)
Integrity of the disco-ligamentous complex	0.304 (0.133-0.483)	0.876 (0.802-0.921)
Neurologic status	1.000 (1.000-1.000)	1.000 (1.000-1.000)
Sum of SLIC and severity scale	0.775 (0.665-0.860)	0.962 (0.942-0.976)

Table 3. The interobserver agreement in SLIC and severity scale among the different observers mean (confidence interval)

	Spine surgeon	Resident of neurosurgery	Neuro-radiologist
Injury morphology	0.927 (0.880-0.956)	0.910 (0.866-0.940)	0.925 (0.885-0.952)
Integrity of the disco-ligamentous complex	0.814 (0.705-0.886)	0.852 (0.782-0.901)	0.963 (0.942-0.976)
Neurologic status	1.000 (1.000-1.000)	1.000 (1.000-1.000)	1.000 (1.000-1.000)
Sum of SLIC and severity scale	0.974 (0.955-0.985)	0.948 (0.921-0.966)	0.963 (0.941-0.977)

SLIC : sub-axial injury classification

Fig. 1. A 33-year-old male was admitted to the hospital due to traffic accident. He complained of C6 dermatome radiculopathy. Cervical spine computed tomography reveals a fracture involving right facet joint at level of C6, 7. Cervical MRI shows the enhancing lesion suggesting disco-ligamentous injury at level of C6, 7. Spine surgeon gave 2 points, resident and radiologist gave 1 point for the patient disco-ligamentous complex injury. The patient's total score was 4 form resident and radiologist, 5 from spine surgeon. Cervical anterior fusion was performed at level of C6, 7.

rized in Table 2. The interobserver agreement was checked differently by the axis of SLIC and severity scale. The interobserver agreement of the injury morphology and total sum of SLIC and severity scale resulted substantial agreement [ICC value=0.603 (CI 0.455-0.731) and ICC value=0.775 (CI 0.665-0.860)]. But, the interobserver agreement of the integrity of the DLC showed fair agreement [ICC value=0.304 (CI 0.133-0.483)]. The intraobserver reliability in SLIC and severity scale was presented as almost perfect agreement. Each Cohen's kappa of the intraobserver reliability in SLIC and severity scale was 0.921 in the injury morphology, 0.876 in the integrity of the DLC, and 0.962 in total sum of SLIC and severity scale. The Cohen's kappa of the neurological status in the interobserver agreement and the intraobserver reliabilty was checked as perfect agreement (ICC value=1.000), and this was resulted by the retrospective review of the medical records.

The more detailed results of the interobserver agreement in SLIC and severity scale among the different observers were shown in Table 3. All interobserver agreement of SLIC and severity scale showed almost perfect agreement in ICC value. In the axis of the injury morphology, the neurologic status and total sum of SLIC and severity scale were similar between observers. But, the agreement of the integrity of the DLC is slightly differently checked : the agreement of neuro-radiologist was higher than the agreement of the spine surgeon or the resident of neurosurgery (ICC value=0.963 vs. ICC value=0.814 or 0.852). Sufficient reproducibility of neuro-radiologist was better than others in reviewing of the integrity of the DLC.

DISCUSSION

The sub-axial spine injuries accounts for the majority of cervical injuries, making up about 65% of fractures and more than 75% of all dislocations[13]. Treatment options are conservative treatment and surgical treatment. Posterior stabilization of the cervical spine is a common surgical procedure which is used in a variety of spinal disorders, including cervical spondylosis, postsurgical deformity or instability, tumor and trauma[4,12]. But, surgical treatment might result in complications, it is not crucial to determine if surgical treatment is necessary. For this reason, many studies reported the classification of spine injuries according to the severity[1,5,8,11].

Historically, Holdsworth[6] first emphasized the importance of the posterior ligamentous complex for maintaining spinal stability in 1970. He published the incipient paradigm for categorizing spinal injuries, which were predicated upon the presumed mechanism of injury. Despite this bold attempt, this system has some limitation of not differentiating the cervical conditions and those affecting the thoracolumbar spine.

In 1982, Allen et al.[1] subsequently proposed their mechanistic classification scheme for subaxial injuries, which was also based upon the findings of plain radiographs. In 1986, Harris et al.[5] took into account the rotational vectors of the forces applied to the spine. But, these systems have some limitations by reviewing only plain radiograph to guess the mechanism of injuries[14]. In this long-standing effort to characterize subaxial trauma, unfortunately no single system has gained widespread acceptance.

In 2007, Subcommittee of the Spine Trauma Study Group introduced the SLIC and severity scale with participation of about 30 spine surgeons. This scoring system consists of injury mor-

phology, integrity of the DLC, and the patient's neurological status. It is for compensating the limitation of previous classification system. The study represents SLIC and severity scale has superior validity and reliability compared to previous study[10].

However, the SLIC and severity scale has some limitations. First, the preliminary study introduced SLIC and severity scale was quite restricted involving only eleven cases. But, related research since the first study collected numerous date and verified the reliability and validity of the SLIC and severity scale, and this research is another support one[2,3]. Secondly, preliminary study participants were limited to only the clinical spine surgeons. For the ideal classification system, available tools are required to communicate not only between spine surgeons but also between neuro-radiologists and residents of neurosurgery who are cooperated with them.

Therefore, in this study the intraobserver reliability and interobserver agreement have been evaluated for three different types of physicians including a spine surgeon, a resident of neurosurgery and a neuro-radiologist. Consequently, the intraobserver reliability and interobserver agreement of SLIC and severity scale are sufficient for classifying cervical spine injury and communicable among physicians of several departments. The Cohen's kappa value usually used for clinical reliability, and at least 0.55 of kappa indicate that valuable reliability[8]. Although Oner et al.[6] argued that this 0.55 standard is too strict to apply to a classification system for traumatic vertebral fracture, but our research showed the excellent of reliability.

In detail, intraobserver reliability among all of three observers showed almost perfect agreement (ICC value >0.8). Also, the interobserver agreement of total score also showed substantial agreement (ICC value=0.775). It is notable interesting that the intraobserver reliability of DLC component was higher in the neuro-radiologist than in the spine surgeon or in the resident of neurosurgery. This may be because of the difficulty in classifying the DLC component to 'intermediate' or 'disrupted', when assessed only by an image study. Consequently, it may be difficult to make objective and definite standard to assess the severity of injury based on the image (Fig. 1).

Nevertheless, this study showed sufficient reproducibility of SLIC and severity scale but also this study has some limitation. First, by using the data of the postoperated patients, the SLIC and severity scale was higher than other previous study (median 7.86, range from 4 to 10). For example, Vaccaro et al.[10] reported 0.83 ICC value in intraobserver reliability and 0.71 ICC value in interobserver agreement which is lower than that of our study, 0.962 and 0.775, respectively. The severely injured patients with distinct imaging results who were participated in the study contributed to high reliability results. Secondly, the high reliability results could affected by the environment that all three physicians worked in same medical center. Another limitation of this study was a data collection by retrospectively, so all neurological status components was accessed reviewing the medical chart, so the ICC value results in perfect agreement.

The prospective study for cervical trauma should be done to overcome these limitations.

CONCLUSION

The SLIC and severity scale is comprehensive and easily applicable tool in spine injured patient. The SLIC and severity scale is suspected as a very useful tool to communicate among spine surgeons, residents of neurosurgery and neuro-radiologists with sufficient reproducibility.

References

1. Allen BL Jr, Ferguson RL, Lehmann TR, O'Brien RP : A mechanistic classification of closed, indirect fractures and dislocations of the lower cervical spine. **Spine (Phila Pa 1976) 7** : 1-27, 1982
2. Anderson PA, Moore TA, Davis KW, Molinari RW, Resnick DK, Vaccaro AR, et al. : Cervical spine injury severity score. Assessment of reliability. **J Bone Joint Surg Am 89** : 1057-1065, 2007
3. Bono CM, Schoenfeld A, Gupta G, Harrop JS, Anderson P, Patel AA, et al. : Reliability and reproducibility of subaxial cervical injury description system : a standardized nomenclature schema. **Spine (Phila Pa 1976) 36** : E1140-E1144, 2011
4. Dvorak MF, Fisher CG, Fehlings MG, Rampersaud YR, Oner FC, Aarabi B, et al. : The surgical approach to subaxial cervical spine injuries : an evidence-based algorithm based on the SLIC classification system. **Spine (Phila Pa 1976) 32** : 2620-2629, 2007
5. Harris JH Jr, Edeiken-Monroe B, Kopaniky DR : A practical classification of acute cervical spine injuries. **Orthop Clin North Am 17** : 15-30, 1986
6. Holdsworth F : Fractures, dislocations, and fracture-dislocations of the spine. **J Bone Joint Surg Am 52** : 1534-1551, 1970
7. Hwang IC, Kang DH, Han JW, Park IS, Lee CH, Park SY : Clinical experiences and usefulness of cervical posterior stabilization with polyaxial screw-rod system. **J Korean Neurosurg Soc 42** : 311-316, 2007
8. Oner FC, van Gils AP, Dhert WJ, Verbout AJ : MRI findings of thoracolumbar spine fractures : a categorisation based on MRI examinations of 100 fractures. **Skeletal Radiol 28** : 433-443, 1999
9. Sanders RW : The problem with apples and oranges. **J Orthop Trauma 11** : 465-466, 1997
10. Vaccaro AR, Hulbert RJ, Patel AA, Fisher C, Dvorak M, Lehman RA Jr, et al. : The subaxial cervical spine injury classification system : a novel approach to recognize the importance of morphology, neurology, and integrity of the disco-ligamentous complex. **Spine (Phila Pa 1976) 32** : 2365-2374, 2007
11. Vaccaro AR, Lehman RA Jr, Hurlbert RJ, Anderson PA, Harris M, Hedlund R, et al. : A new classification of thoracolumbar injuries : the importance of injury morphology, the integrity of the posterior ligamentous complex, and neurologic status. **Spine (Phila Pa 1976) 30** : 2325-2333, 2005
12. Viera AJ, Garrett JM : Understanding interobserver agreement : the kappa statistic. **Fam Med 37** : 360-363, 2005
13. Watson-Jones R : The results of postural reduction of fractures of the spine. **J Bone Joint Surg Am 20** : 567-586, 1938
14. Whang PG, Patel AA, Vaccaro AR : The development and evaluation of the subaxial injury classification scoring system for cervical spine trauma. **Clin Orthop Relat Res 469** : 723-731, 2011