



# Neurocirugía



<https://www.revistaneurocirugia.com>

## C0086 - INTEROBSERVER AGREEMENT OF YOUNG NEUROSURGEONS AND TEACHABILITY FOR SUBAXIAL CERVICAL SPINE INJURY CLASSIFICATION SYSTEMS

A. Grin, I. Lvov, A. Talygov, A. Kordonskiy, U. Khushnazarov and V. Krylov

*Sklifosovsky Research Institute of Emergency Care, Moscow, Russia.*

### Resumen

**Objectives:** To evaluate interobserver agreement and neurosurgeon teachability for Allen-Ferguson, Harris, Argenson, AOSpine, SLIC and CSISS classifications.

**Methods:** This work was part of a multicentre study of lower cervical spine injury classification agreement. The participants included 11 neurosurgical residents and beginning spine surgeons. Participants received the author's booklet illustrating in detail each classification, an answer form, and a flash drive with de-identified CT, MRI, and neurological data for 64 consecutive patients, operated on between 2013 and 2017. All raters assessed each case twice within a 6-8 week interval. The main group (6 raters) participated in seminars on classification analysis in between assessments. The control group (5 raters) did not participate in the seminars. Interobserver reliability was assessed using Fleiss' kappa (K) and ICC. Teachability was measured as the Fleiss' kappa or ICC difference between two assessments (delta kappa or delta ICC).

**Results:** In the main group, the AOSpine and Argenson morphological classifications demonstrated the best teachability (0.38 and 0.35 respectively) with substantial agreement (K = 0.62 and 0.61 respectively). Interobserver agreement for Allen-Ferguson and Harris schemes was moderate and fair (K = 0.33 and 0.4 respectively) and teachability for both scales was < 0.1. The SLIC demonstrated the best numerical scale teachability and interobserver ICC (Delta ICC 0.18, delta kappa 0.13). The correlation was high (0.73), but answer congruence was only fair (interobserver K = 0.22). Teachability was the lowest for CSISS (< 0.006) with substantial correlation (0.61) and low answer congruence (K = 0.14). In the control group, delta kappa for all scales were less than 0.06.

**Conclusions:** Argenson, AOSpine and SLIC scales may be preferable for clinical usage. Short learning courses provided significant increases in interobserver agreement for these scales. Further comparative studies of existing classifications should be performed to determine the best scheme for day-to-day practice.