The Bone & Joint Journal

Hosman AJF, Barbagallo G, the SCI-POEM Study Group, van Middendorp JJ. Neurological recovery after early versus delayed surgical decompression for acute traumatic spinal cord injury. *Bone Joint J*. 2023;105-B(4):400-411.

https://doi.org/10.1302/0301-620X.105B4.BJJ-2022-0947.R2

14 April 2023

Sir,

We read the recent article by Hosman et al[1] with great interest.

The impact of surgical intervention within 12 hours of injury on the recovery of lower extremity motor score (LEMS) at 12 months was the main focus of the study, which is at odds with current guidelines that define early intervention as being within 24 hours.

The study showed that patients who underwent early surgical decompression (within 12 hours) improved 4.3 LEMS more (95% CI -0.3 to 8.8; p = 0.065) than patients who underwent late surgery (> 12hours < 12days). Following adjusted analysis with multiple imputation, the disparity between both groups was 2.2 (95% CI -1.5 to 5.9), thereby showing an insignificant effect of early surgery on LEMS recovery.

The authors are to be commended for their substantial effort in initiating a large European prospective multicentre cohort study on this crucial subject but we must raise some concerns on their interpretation of its findings.

When the study was conceived, more than ten years ago, there was still 'equipoise' among spinal trauma specialists on the issue of timing of surgery in traumatic spinal cord injury (tSCI). The study was designed on the assumption that there were variations in practice between different European centres, and that a prospective cohort study would yield comparable groups of patients who were treated differently in terms of timing. However, it seems that by the time the study started recruiting patients, most of the participating centres had adopted a policy of early intervention, and they seem to have been quite successful in doing so since 54.1% of all patients underwent surgery within 12 hours. While we should congratulate our European colleagues for this achievement, it may have undermined the basic design of the SCI-POEM study, as this introduced a strong selection bias in that all centres involved seem to have done their best to treat patients as soon as possible. 'Late' cases in that light can be seen as a 'failure to follow a consistent policy on tSCI'. Why these 'late' cases didn't undergo early surgery could probably be attributed to other unidentified factors. It may also have caused the substantial differences between the early and late cohorts at baseline. These specific and significant differences at baseline between the two cohorts are our primary concern about the conclusions drawn from this study.

Patients in the early cohort were significantly younger (44.7 years versus 49.1 years), had more severe injuries (Injury Severity Score (ISS) 19.1 versus 15.6) and more severe neurological injuries (47.1% Abbreviated Injury Scale (AIS) A), and more often involved the thoracic spine. By contrast, the late cohort consisted of older patients with less severe neurological injuries (43.8% AIS D) involving mostly the cervical spine. The authors attempted to correct for this significant baseline imbalance by carrying out propensity score adjustments, but despite these efforts, the baseline AIS grade remained a persisting confounding factor. In their unadjusted analysis they observed that patients treated within 12 hours gained 4.3 more LEMS than patients treated later. This is an interesting finding, as the early cohort included more severe neurological injuries, which typically have an adverse impact on potential LEMS recovery.

The authors selected LEMS as their outcome measure because recovery in motor function may not always result in AIS conversion: even if AIS conversion occurs, it may not always lead to any meaningful functional recovery. However, the authors did not acknowledge that the recovery of LEMS is highly variable depending on the initial severity of the neurological injury, and that recovering upper extremity motor function may be more important for individuals with severe cervical injuries. When the authors carried out a subgroup analysis on AIS grade, the positive impact of early surgery on LEMS recovery remained for patients with a baseline AIS A (4.6 LEMS) and AIS B (4.2 LEMS), whereas for AIS C and D patients the difference in LEMS recovery did not differ much between the groups (-1.3 and 1.0 LEMS). Moreover, patients with tetraplegia in the early cohort had significantly lower LEMS at baseline (10.5 versus 25.3) and experienced 16.7 LEMS recovery versus 14.0 in the late cohort. While this difference in LEMS might seem trivial, the recovery of 16.7 LEMS for a patient with 10.0 LEMS at baseline could potentially mean a shift from non-functional to functional recovery.

Our second concern is the high rate of missing data (31% early cohort, 36% late cohort) on the primary outcome (LEMS at 12 months). Remarkably, 29% of all AIS A patients in the early cohort and 44.7% of the AIS A patients in the late cohort had missing data on LEMS at follow-up which was imputed by using a predictive mean matching method. Moreover, patients in the late cohort who had missing data at 12 months had a significantly lower LEMS (3.5 versus 26.0) at baseline than patients who did complete this follow-up. While predictive mean matching can be used to impute missing data by predicting values using a regression model and selecting individuals with similar predicted values for the covariate of interest, it may be difficult to predict missing data accurately if a confounding factor remains imbalanced, such as baseline AIS and baseline LEMS. Therefore, caution should be exercised when using predictive mean matching in this situation, since this limits the validity of the imputed data.

Another concern is the large number of variables used for the propensity score model. To match patients accurately for all the selected variables it is preferable to have sufficient patients with overlapping baseline characteristics. Propensity score matching requires sufficient observations to increase the probability of finding proper matches: because the ratio between the number of parameters and observations is high, it is unlikely to succeed. This is indeed seen by the failure to remove the significant difference in baseline AIS between the groups. In general, when including more variables to any model, the bias of the estimated effect of surgical timing may decrease. However, due to the small number of patients, the variance of the estimated effect will also increase. This in turn will decrease the precision of the estimated effect of surgical timing, yielding non-significant results in the model.

Finally, the authors compared patients who underwent stabilization alone with those who underwent actual decompression of the injured spinal cord. While the rate of laminectomy did not differ between the two groups, it remains unclear whether the spinal cord was actually

decompressed in either group since pre- and postoperative MRIs are not routinely performed. Currently, more and more studies show that the actual extent of decompression of the spinal cord might be more relevant for prevention of secondary neurological injury than was previously recognized.[2-5] Interestingly, patients with more severe SCI will have profound swelling of the spinal cord, which can progress over the following days.[5,6]

The SCI-POEM group should be congratulated on this excellent study, but they should admit that they failed in one of its primary aims, namely to create similar groups of patients treated in different timeframes. Instead, they could have compared the recovery patterns of this cohort which has excellent follow-up with similar historical cohorts such as the Surgical Timing in Acute Spinal Cord Injury Study (STASCIS).[7]

P. V. ter Wengel, MD, PhD,
R. Martin,
Department of Neurosurgery, Haaglanden Medical Center,
the Hague, The Netherlands
W. P. Vandertop, MD, PhD,
Department of Neurosurgery, Amsterdam University Medical Center,
Amsterdam, The Netherlands
F. C. Öner, MD, PhD,
Department of Orthopedic Surgery, University Medical Center Utrecht,
Utrecht, The Netherlands.

1. **Hosman AJF, Barbagallo G, the SCI-POEM Study Group, van Middendorp JJ.** Neurological recovery after early versus delayed surgical decompression for acute traumatic spinal cord injury. *Bone Joint J*. 2023;105-B(4):400-411.
2. **Aarabi B, Olexa J, Chryssikos T, et al**. Extent of Spinal Cord Decompression in Motor Complete (American Spinal Injury Association Impairment Scale Grades A and B) Traumatic Spinal Cord Injury Patients: Post-Operative Magnetic Resonance Imaging Analysis of Standard Operative Approaches. *J Neurotrauma*. 2019;36(6):862–876.
3. **Phang, I, Werndle MC, Saadoun S, et al**. Expansion duroplasty improves intraspinal pressure, spinal cord perfusion pressure, and vascular pressure reactivity index in patients with traumatic spinal cord injury: injured spinal cord pressure evaluation study. *J Neurotrauma*. 2015;32(12):865–874.
4. **Hogg FRA, Kearney S, Zoumprouli A, Papadopoulos MC, Saadoun S**. Acute Spinal Cord Injury: Correlations and Causal Relations Between Intraspinal Pressure, Spinal Cord Perfusion Pressure, Lactate-to-Pyruvate Ratio, and Limb Power. *Neurocrit Care*. 2021;34(1):121–129.
5. **Aarabi B, Simard JM, Kufera JA, et al**. Intramedullary lesion expansion on magnetic resonance imaging in patients with motor complete cervical spinal cord injury. *J Neurosurg Spine*. 2012;17(3):243–250.
6. **Le E, Aarabi B, Hersh DS, et al**. Predictors of intramedullary lesion expansion rate on MR images of patients with subaxial spinal cord injury. *J Neurosurg Spine*. 2015;22(6):611–621.
7. **Fehlings MG, Vaccaro A, Wilson JR, et al.** Early versus delayed decompression for traumatic cervical spinal cord injury: results of the Surgical Timing in Acute Spinal Cord Injury Study (STASCIS). *PLoS One*. 2012;7(2):e32037.

Conflict of Interest: Prof. F.C. Öner was involved in the SCI-POEM group and participated in the study. All other authors report no conflict of interest and have no financial disclosures.