**APPENDICES: SUPPLEMENTAL MATERIAL**

**These appendices contain detailed abstraction and results tables as well as documentation of the systematic literature search and assessment of study quality. A list of studies excluded at full-text level is included at the end.**

**Key question 1: Demographics**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Author** | **Study Design** | **Population** | **Inclusion/ Exclusion Criteria** | **Baseline Population Definition and Characteristics** | **Treatment Groups** | **Baseline MRI Characteristics**  **and Assessment** | **MRI Timing: Injury and Intervention** |
| Papadopou los (2002) | Prospective cohort | *Protocol group:*  N = 66  Sex: 68.2% (45/66) male  Mean age: 32 ± 2.1 (2-92) years  Mean F/U time: 31.6 ± 3.6 (2-92) months  F/U %: NR  *Reference group:*  N = 25  Sex: 76.0% (19/25) male  Mean age: 42 ± 4.8 (2-89) years  Mean F/U time: 36.2 ± 5.7 (2-96) onths  F/U %: NR | *Inclusion:*   * Clinically and radiographically confirmed acute traumatic spinal cord and vertebral column injury   *Exclusion:*   * Central cord syndrome secondary to cervical spondylotic stenosis but without acute spinal column injury * Received definitive surgical treatment outside the University of Michigan Medical Center | *Population:* 91 consecutive patients suffering from traumatic closed cervical (C1-T1) SCI who received acute care at the University of Michigan Medical Center between 1990 and 1997. Patients were managed by a treatment protocol (protocol group) or were part of a reference group.  *Admitting Frankel score*  Protocol group   * Frankel Grade A: 57.6% (38/66) * Frankel Grade B: 18.2% (12/66) * Frankel Grade C: 13.6% (9/66) * Frankel Grade D: 10.6% (7/66)   Reference group   * Frankel Grade A: 64.0% (16/25) * Frankel Grade B: 16.0% (4/25) * Frankel Grade C: 0.0% (0/25) * Frankel Grade D: 20.0% (5/25)   *Injury Severity Score:*   * Protocol: 28 ± 1.6 (9-75) * Reference: 35 ± 5.5 (14-75) | * Protocol group (n=66):   Patients with cervical spinal malalignment causing suspected spinal canal compromise were placed in skeletal traction. Once optimal spinal alignment was achieved, patients underwent MRI. Patients with MRI-documented SCC underwent emergency surgical decompression and spinal column stabilization. Patients without SCC were treated with either surgery for internal spinal column stabilization within 24-48 hours or placed in a definitive external orthosis.  *Surgical Intervention*:   * Discectomy/fusion: 16.7% (11/66) * Corpectomy/fusion: 21.2% (14/66) * Posterior fusion: 34.8% (23/66) * Laminectomy/fusion: 7.6% (5/66) * Combined anterior/posterior fusion: 4.5% (3/66) * Alignment/external orthosis: 15.2% (10/66) * Reference group (n=25):   Reference group patients were treated outside of the protocol because of contraindication to MRI, the need for an emergent surgical procedures, or admitting surgeon preferences | *MRI characteristics:*  T1- and T2- weighted MR imaging of the cervical spine was performed using a 1.5-Tesla scanner (protocol group only)  *MRI assessment:* NR | *MRI timing with regard to injury:* Mean time from injury to MRI was 7.7 ± 0.4 hours\* (for protocol group only)  *MRI timing with regard to intervention:* After MRIs were obtained, protocol group patients underwent emergency surgical intervention or non-emergent surgery if indicated within 24-48 hours. Mean time from injury to operative decompression was 12.6 ± 1.3 hours\*. |

F/U: follow-up; MRI: magnetic resonance imaging; NR: not reported; SCC = spinal cord compression; SCI = spinal cord injury

\* Reported as mean ± standard error

**Key Question 1: Results and Outcomes**

|  |  |  |  |
| --- | --- | --- | --- |
| **Author** | **Treatment Groups and Potentially Confounding Factors Evaluated** | **Outcome Measures Evaluated\*** | **Significant Results**† |
| ***Influence of baseline MRI characteristics upon treatment strategy(ies) and neurologic, functional, and safety outcomes*** | | | |
| Papadopoulos (2002) | *Treatment group*   1. Protocol group (n = 66) 2. Reference group (n = 25)   *Potentially confounding factors*  Demographic   * Age * Sex   Clinical   * Mechanism of injury * Admitting injury level * Admitting Injury Severity Score * Admitting Frankel grade | *Neurologic*   * Frankel grade at last follow-up | Effect of treatment group on Frankel grade improvement at last follow-up:   * Protocol group (*P < .006*): Patients in the Protocol group improved 7/10 of 1 Frankel grade more than patients in the Referene group |

\* Only reported outcome measures that related to study question

† p<.05; significant results as reported by the authors based upon results of multivariate regression analysis

**Key Question 2: Demographics**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Author** | **Study Design** | **Population** | **Baseline Inclusion/ Exclusion Criteria** | **Baseline Population Definition and Characteristics** | **Baseline MRI Characteristics**  **and Assessment** | **MRI Timing: Injury and Intervention** |
| Aarabi (2011) | Retrospective cohort | N = 42  Sex: 83.3% (35/42) male  Mean age: 58.3 (32-87) years  Mean F/U time: 29.1 ± 27 months  F/U %: 85.7% (42/49) | *Inclusion:*   * ≥ 18 years old * History of trauma * Acute traumatic CCS\* * Admission ASIA motor score < 96 * Evidence of spinal stenosis on imaging studies w/ or w/o distractive extension Stage 1 * Underwent surgical decompression of spinal cord * ASIA Impairment Scale Grades C or D   *Exclusion:*   * Fracture dislocations including distractive extension Stage 2 * SCI w/o cord compression on MRI * Acute disc prolapse * Admission ASIA motor scores of 96-100 * Patients with TBI or baseline dementia * ASIA Impairment Scale Grades A, B or E | *Population:* Patients with acute traumatic CCS who were admitted to the R. Adams Cowley Shock Trauma Center from January 1, 2000 to April 30, 2008  *Intervention:* Patients chose early or late (w/in 6 weeks of injury) decompression surgery; surgical technique was influenced by imaging studies   * Anterior cervical discectomy or corpectomy and fusion: 40.5% (17/42) * Laminectomy or laminoplasty ± fusion: 33.3% (14/42) * ACDF or corpectomy and laminectomy ± fusion: 26.2% (11/42)   *SCI Severity (ASIA Impairment scale)*   * ASIA grade C: 21.4% (9/42) * ASIA grade D: 78.6% (33/42)   *Mean injury severity score: 16.8*  *Mean admission ASIA motor scale score: 63.8*  *Mean total sensory score: 190* | *MRI characteristics:* MR images (40/42 patients) were obtained on 41 of 42 patients; films on another patient were corrupt and could not be retrieved. T2-weighted images were used for image assessment. CT scans of the cervical spine were obtained for all patients with coronal, sagittal and axial reformatting.  *MRI assessment:* All images were measured independently by two radiologists and the senior author in a blinded fashion. | *MRI timing with regard to injury:* Imaging studies were obtained following stabilization. MR images were obtained, on average, 5.8 hours after admission, and CT images were obtained, on average, 2.3 hours after admission.  *MRI timing with regard to intervention:* MRI (41/42 patients) and CT scans (42/42 patients) were obtained prior to surgery; timing NR. Median days after injury to operation was 2.8 days. |
| Boldin (2006) | Prospective cohort | N = 29  Sex: 65.5% (19/29) male  Mean age: 43.5 ± 18.1 (18-86) years  Median F/U time: 35 (24-65) months  F/U %: NR | *NR* | *Population:* Patients with closed cervical SCI and neurologic deficit who were admitted to University Clinic of Traumatology of Graz from September 1996 through September 2001  *Intervention:* All patients were treated operatively within 2-9 (mean: 4.3) hours after injury by closed or open reduction and ventral stabilization.  *SCI Severity (ASIA Impairment scale)*   * ASIA grade A: 27.6% (8/29) * ASIA grade B: 37.9% (11/29) * ASIA grade C: 27.6% (8/29) * ASIA grade D: 6.9% (2/29) * ASIA grade E: 0% (0/29) | *MRI characteristics:* T1- and T2- weighted MR images were obtained using a 1.5 Tesla superconducting MR scanner with a surface coil, a 3-mm slice thickness for T1 and 2-mm thickness for T2 images.  *MRI assessment:* NR. | *MRI timing with regard to injury and intervention:* MR was performed within 5-12 (median: 8) days after injury. MR imaging was obtained after intervention. |
| Flanders (1996) | Retrospective cohort | N = 104  Sex: 87.5% (91/104) male  Mean age: 34 (17-70) years  Mean F/U time: ≥ 12 months  F/U %: 58.1% (104/179) | *Inclusion:*   * Complete initial clinical assessment of motor power ≤ 1 week after injury, final assessment of motor power ≥ 12 months after injury * Single diagnostic-quality MR study of the cervical spine performed ≤ 72 hours from initial injury * Injuries resulting from non-penetrating trauma * Isolated cervical SCIs from neurologic levels C3 to T1   *Exclusion:* NR | *Population:* Patients with cervical SCI who were admitted to the Regional Spinal Cord Injury Center of the Delaware Valley from July 1988 through July 1993  *Intervention:* Closed reduction, subluxation, or both  *SCI Severity (ASIA Impairment scale)*   * ASIA grade A: 41.3% (43/104) * ASIA grade B: 22.1% (23/104) * ASIA grade C: 26.9% (28/104) * ASIA grade D: 9.6% (10/104) * ASIA grade E: 0% (0/104) | *MRI characteristics:*  Sagittal T1- and T2- weighted MR images were acquired using a 1.5 Tesla superconducting unit to obtain sagittal spin-echo sequences  *MRI assessment:* All images were given a quality rating of satisfactory† or incomplete; if unsatisfactory or incomplete, patients were excluded from analyses. All spinal cord trauma locations were recorded by a single neuroradiologist who was blinded to the clinical status of each subject | *MRI timing with regard to injury and intervention:* MRI was performed within 72 hours of injury before or after surgical stabilization. |
| Miyanji (2007) | Prospective cohort | N = 100  Sex: 79.0% (79/100) male  Mean age: 45 (17-96) years  Mean F/U time: 7.3 (1-35) months  F/U %: NR | *Exclusion:*   * Concomitant head injury and Glasgow coma scale score <15 | *Population:* Consecutive patients with cervical spine trauma who were admitted to the Toronto Western Hospital or University of Maryland Medical System from March 2000 through March 2005  *Intervention:* NR  *SCI Severity (ASIA Impairment scale)*   * Complete (ASIA grade A): 26% (26/100) * Incomplete (ASIA grade B, C, D): 51% (51/100) * Normal (ASIA grade E): 22% (22/100) * Unknown: 1% (1/100) | *MRI characteristics:*  Sagittal T1- and T2- weighted MR images were obtained using a 1.5-Tesla magnet and standardized MR imaging protocols for the acutely injured.  *MRI assessment:* All images were analyzed by a spine fellow board-certified in orthopedic surgery with 7 yrs experience; this individual was blinded to the patients’ clinical and neurologic data. | *MRI timing with regard to injury:* MR images were obtained within 24-48 (median: 24) hours of injury.  *MRI timing with regard to intervention:* NR |
| Selden (1999) | Retrospective cohort | N = 55  Sex: 65.5% (36/55) male  Mean age: 29.2 (2-92) years  Mean F/U time: 18.5 (1-74) months  F/U: NR | *Exclusion:*   * Patients with contraindications to scanning * Patients needing other emergency surgical procedures * Inadequate scans (obtained on a 0.3-Tesla scanner) or missing scans * Patients with CCS secondary to cervical spondylitic stenosis, but w/o evidence of acute SCI | *Population:* Patients with traumatic closed cervical (C1-T1) myelopathy admitted to the University of Michigan Medical Center from May 1990 through December 1996.  *Intervention:* Patients with MRI-documented spinal cord compression underwent immediate surgical decompression and fusion. Patients without spinal cord compression underwent urgent surgery for spinal column stabilization within 48 hours.   * Discectomy/fusion: 12.7% (7/55) * Corpectomy/fusion: 25.5% (14/55) * Posterior fusion: 40.0% (22/55) * Laminectomy/fusion: 5.5% (3/55) * Combined anterior/posterior fusion: 1.8% (1/55) * Reduction/external orthosis only: 14.5% (8/55)   *SCI Severity (Initial Frankel Grade)*   * Frankel grade A: 58.2% (32/55) * Frankel grade B: 16.4% (9/55) * Frankel grade C: 14.5% (8/55) * Frankel grade D: 10.9% (6/55) | *MRI characteristics:*  Sagittal and targeted axial T1- and T2- weighted MR imaging examinations were performed using a 1.5-Tesla scanner.  *MRI assessment:* All images were interpreted by an attending neuroradiologist who was blinded to the clinical status of the patients. | *MRI timing with regard to injury:* MR images were obtained within 21 (mean 7.8) hours of injury.  *MRI timing with regard to intervention:* All MRIs were obtained prior to intervention. Emergency interventions occurred at a mean of 13.6 hours after injury, and urgent interventions occurred at a mean of 28.7 hours after injury. |
| Shepard (1999) | Prospective cohort (secondary analysis of data collected prospectively in a RCT) | N = 191  Sex: 84.8% (162/191) male  Age (years)   * 14-22: 25.6% (49/191) * 23-34: 28.3% (54/191) * 35-50: 21.5% (41/191) * > 50: 24.6% (47/191)   Mean F/U time: 6 weeks  F/U %: NR | *Inclusion:*   * MRI performed within 72 hours of injury | *Population:* Patients randomized in the National Acute Spinal Cord Injury Study (NASCIS) 3 between December 1991 and September 1995 who had MR imaging within 72 hours of injury. The purpose of the RCT was to evaluate the efficacy and safety of methylprednisolone infused over 48 hours, or tirilazad mesylate administered every 6 hours for 48 hours with 24 hours of methylprednisolone.  *Intervention:* Spinal surgery was performed within 3 days in 45.0% (86/191) of patients.  *Extent of Injury (determined by neurologic and radiologic exams)*   * Complete: 39.3% (75/191) * Incomplete: 45.5% (87/191) * Normal: 15.2% (29/191) | *MRI characteristics:* Sagittal and axial T1- and T2- weighted MR imaging with 3-5 mm slices. Technical information on the MR imaging unit was not ascertained.  *MRI assessment:* Results of MR images were transcribed from the MRI report onto the NASCIS forms by study neurosurgical nurses, or were read directly from the image onto the NASCIS forms by the neurosurgeon associated with NASCIS 3. | *MRI timing with regard to injury:* MR images were obtained within 72 hours of injury. 51.8% (99/191) received MRI within 9 hours of injury, and 48.2% (92/191) received MRI > 9 hours after injury.  *MRI timing with regard to intervention:* NR |
| Wilson (2012) | Prospective cohort (combined two prospectively collected databases with similar data elements) | N = 376  Sex: 78.2% (294/376) male  Mean age: 43.2 ± 16.9 years  F/U time   * 6 months: 17.6% (66/376) * 12 months: 82.4% (310/376)   F/U %: 54.0% (376/696) | *Inclusion:*   * ≥ 16 years old * Traumatic SCI and neurologic deficit (AISA impairment scale grade A-D) * ASIA neurologic examination performed within 3 days of injury   *Exclusion:*   * Patients with severe head injury | *Population:* Patients who were prospectively enrolled in the North American Clinical Trials Network for SCI registry database or the Surgical Timing in Acute Spinal Cord Injury Study multicenter cohort study (August 2002 to September 2009).  *Intervention:* Decisions surrounding surgery, including timing, were made by the attending spine surgeon in each case. Mean time to decompressive surgery post-injury was 76.1 (± 338.9) hours.  *SCI Severity (ASIA Impairment scale)*   * ASIA grade A: 36.2% (136/376) * ASIA grade B: 16.8% (63/376) * ASIA grade C: 15.4% (58/376) * ASIA grade D: 31.7% (119/376)   *Initial ASIA motor score > 50:* 32.9% (124/376) | *MRI characteristics:* T1- and T2- weighted images were obtained using standard protocols.  *MRI assessment:* Images were read by a site-specific neuroradiologist. | *MRI timing with regard to injury:* MR images were obtained within 3 days of injury.  *MRI timing with regard to intervention:* NR |

ACDF = anterior cervical decompression and fusion; ASIA = American Spinal Injury Association; CCS = central cord syndrome; CT = computed tomography; F/U = follow up; MR = magnetic resonance; MRI = magnetic resonance imaging; NASCIS = National Acute Spinal Cord Injury Study; NR: not reported; SCI = spinal cord injury; TBI: traumatic brain injury

\* Aarabi 2011: Central cord syndrome was defined as incomplete spinal cord injury with weaker upper extremities than lower extremities

† Flanders 1996: Satisfactory quality rating was defined on a sagittal T1-weighted spin echo (or the combination of sagittal T2-weighted fast spin-echo acquisition and sagittal gradient-echo) sequence

**Key Question 2: Results and Outcomes**

|  |  |  |  |
| --- | --- | --- | --- |
| **Author** | **Prognostic Factors Evaluated** | **Outcome Measures Evaluated\*** | **Significant Results**† |
| ***Spinal cord lesion characteristics, pattern and length identified on baseline MRI as predictors of neurologic, functional, and safety outcomes*** | | | |
| Aarabi (2011) | *Radiographic*   * Percentage of MCC‡ * Percentage of MSCC‡ * Sagittal diameter of stenotic spinal canal at the point of MSCC (mm) * Length of parenchymal damage (signal change on MRI; mm)   *Secondary (potentially confounding) factors*  Demographic   * Age (older, younger)   Clinical   * Admission ASIA motor score * Mechanism of injury * Number of stenotic skeletal segments (1, 2, ≥3) * Surgical technique (front, back, circumferential decompression) * Time delay from injury to surgery | *Neurologic*   * Follow-up ASIA motor score   *Function and Pain*   * Follow-up FIM score * Manual dexterity level‡ * Dysesthetic pain level‡   *Safety*  NR | Predictors of worse neurologic recovery (lower follow-up ASIA motor score):   * Lower MCC percentage (*P = 0.02*) * Lower sagittal diameter of stenotic spinal canal at the point of MSCC (*P = 0.02*) * Lower admission ASIA motor score (*P = .003*)   Predictors of less independence with activities of daily living (lower follow-up FIM score):   * Lower MCC percentage (*P = 0.02*) * Lower admission ASIA motor score (*P = .03*) * Older patients (*P = .02*)   Predictors of less manual dexterity at latest follow-up:   * Longer parenchymal damage (*P* = 0.002) * Lower admission ASIA motor score (*P = .0002*)   Predictors of greater dysesthetic pain at latest follow-up:   * Longer parenchymal damage (*P = 0.04*) * Younger patients (*P = .02*) |
| Boldin (2006) | *Radiographic*   * Cord edema * Intramedullary hemorrhage * Edema lesion length * Hemorrhage lesion length   *Secondary (potentially confounding) factors*  Clinical   * Baseline ASIA impairment grade | *Neurologic*   * Follow-up ASIA impairment grade   *Function and Pain*  NR  *Safety*  NR | Predictor of retaining complete SCI at follow-up (ASIA grade A):   * Length of edema: Each millimeter increase in edema length was associated with an increased risk of retaining complete SCI at follow-up (OR = 1.15, 95% CI: 1.03-1.29; *P = .017*)   Predictor of retaining complete SCI at follow-up (ASIA grade A), addition of hematoma length to the statistical model:   * Length of hematoma: Each millimeter increase in hematoma length was associated with an increased risk of retaining complete SCI at follow-up (OR = 1.81, 95% CI: 1.09-3.02; *P = .022*) |
| Flanders (1996) | *Radiographic*   * Presence of hemorrhage * Rostral point of edema * Edema lesion length   *Secondary (potentially confounding) factors*  Clinical   * Initial ASIA motor score | *Neurologic*   * Upper extremity motor function§ * Lower extremity motor function§ * Upper extremity muscles with minimally useful function§ * Lower extremity muscles with minimally useful function§   *Function and Pain*  NR  *Safety*  NR | Predictors of decreased upper extremity motor function at least 12 months after injury:   * Presence of hemorrhage (*P = .001*) * Longer edema lesion length (*P = .033*) * Lower initial ASIA motor score (*P = .002*)   Predictors of decreased lower extremity motor function at least 12 months after injury:   * Presence of hemorrhage (*P < .001*) * Longer edema lesion length (*P = 0.026*) * Lower initial ASIA motor score (*P < .001*)   Predictors of decreased upper extremity muscles with useful function at least 12 months after injury:   * Presence of hemorrhage (*P < .001*) * Lower initial ASIA motor score (*P = .048*)   Predictors of decreased lower extremity muscles with useful function at least 12 months after injury:   * Presence of hemorrhage (*P < .001*) * Longer edema lesion length (*P < .001*) * Lower initial ASIA motor score (*P < .001*) |
| Miyanji (2007) | *Radiographic*   * MCC\*\* * MSCC\*\* * Intramedullary hemorrhage * Cord edema * Cord swelling * Soft-tissue injury\*\* * Preinjury stenosis * Disk herniation * SCI lesion length\*\*   *Secondary (potentially confounding) factors*  Clinical   * Baseline ASIA motor score * Baseline ASIA impairment scale | *Neurologic*   * Last follow-up ASIA motor score   *Function and Pain*  NR  *Safety*  NR | Predictors of worse neurologic recovery (lower follow-up ASIA motor score):   * Lower baseline ASIA motor score (*P < .001*) * Intramedullary hemorrhage (*P = .002*) * Cord swelling (*P = .054*) |
| Selden (1999) | *Radiographic*   * Maximal cross-sectional diameter within swollen length of the cord * Spinal cord compression * Length of cord swelling * Length of cord edema * Length of cord hematoma   *Secondary (potentially confounding) factors*  Clinical   * Admission Frankel grade | *Neurologic*   * Follow-up Frankel grade   *Function and Pain*  NR  *Safety*  NR | Predictor of poor neurologic recovery at last follow-up (Frankel grade):   * Longer rostrocaudal length of spinal cord hematoma (*P = 0.028*) * Worse Frankel grade at admission (*P < 0.001*) |
| Shepard (1999) | *Radiographic*   * Cord edema * Hemorrhage * Contusion   *Secondary (potentially confounding) factors*  Clinical   * Drug protocol (randomized group) * Extent of injury (complete/incomplete) * Spinal cord surgery (yes/no) * Baseline pin prick score†† * Baseline light touch score†† * Baseline motor function score†† | *Neurologic*   * Pin prick score†† * Light touch score†† * Motor function score††   *Function and Pain*  NR  *Safety*  NR | Predictors of worse motor function recovery at 6 weeks:   * Evidence of cord edema (*P = .06*)   Predictors of worse pin prick recovery at 6 weeks:   * No potential prognostic factors reached statistical significance   Predictors of worse light touch recovery at 6 weeks:   * Evidence of cord edema (*P = .05*) |
| Wilson (2012) | *Spinal cord pattern*   * MRI signal characteristics consistent with edema or hemorrhage   *Secondary (potentially confounding) factors*  Demographic   * Age   Clinical   * Initial ASIA motor score (≤ 50 or > 50) * Initial ASIA impairment scale grade | *Neurologic*  *Function and Pain*   * FIM motor score‡‡ * Functional independence (dichotomized variable)‡‡   *Safety*  NR | Predictors of worse FIM score at 1 year:   * Initial ASIA motor score ≤ 50 (*P < .01*) * Lower initial ASIA impairment score (*P < .01*) * Older age (*P < .01*)   Predictors of functional dependence at 1 year (logistic regression):   * Initial ASIA motor score ≤ 50 (*P < .01*) * Lower initial ASIA impairment score (*P < .01*) * Older age (*P < .01*) |

ASIA = American Spinal Injury Association; CI = confidence interval; FIM = functional independence measure; MCC = maximum canal compromise; MRI = magnetic resonance imaging; MSCC = maximum spinal cord compression; NR = not reported; OR = odds ratio; SCI = spinal cord injury

\* Only reported outcome measures related to study question

† p<.05; significant results as reported by the authors based upon multivariate regression analyses, adjusted for baseline neurologic status.

‡ Aarabi (2011): *MCC* = {1 - Di/ [1/2(Da + Db)]} × 100, where Di is the midsagittal diameter of the spinal canal at the point of maximum compression, Da is the diameter of the spinal canal one segment above the highest level of spinal stenosis, and Db is the diameter of the spinal canal one segment below the lowest level of spinal stenosis. *MSCC* = {1 - di/ [1/2(da + db)]} × 100, where da is the diameter of the spinal cord at a normal segment above the highest level of spinal stenosis, db is the diameter of the spinal cord one segment below the lowest level of stenosis, and di is the diameter of the spinal cord at the level of maximum stenosis. *Manual dexterity level* was defined as the patient’s own perception of dexterity and skill in the following tasks: 1) using a keyboard, 2) playing a musical instrument, 3) buttoning his/her shirt, 4) grooming, and 5) writing. *Dysesthetic pain level* was a subjective rating on a 0 to 10 analog scale.

§ Flanders (1996): *Extremity motor function* was assessed by testing key individual muscles in each of the upper and lower extremities. Each muscle received a score of 0 to 5 for a total possible motor index score of 50 for upper extremity function and 40 for lower extremity function. *Minimally useful motor function* in an individual muscle was defined as a score of 3 or better on the 5-point manual muscle test. A score of 3 represents the ability for active movement with a full range of motion against gravity.

\*\* Miyanji (2007): *MCC/MSCC* was defined as described by Fehlings MG, Rao SC, Tator CH, et al. The optimal radiologic method for assessing spinal canal compromise and cord compression in patients with cervical spinal cord injury. Part II: Results of a multicenter study. Spine 1999;24 (6):605-613. *Soft-tissue* *injury* was defined as an increased signal intensity of the perivertebral tissues on T2-weighted images. *Spinal cord lesion length* was defined as the distance between the most cephalic and the most caudal extent of the cord signal intensity change on T2-weighted images.

†† Shepard (1999): *Pin prick/light touch score* was determined as follows: 29 spinal cord segments were tested bilaterally for response to pin prick and light touch, each of which were scored at 1=absent, 2=dysfunctional and 3=normal. Sensory scores ranged from a total score of 87 (normal response) to a score of 29 (no response in any segment). Responses for the right side were summed, unless some cord segments could not be tested, in which case responses for the left side were used. *Motor function score* was determined by bilaterally measuring fourteen muscle roots, with 0=no contraction, 1=flicker/trace of contraction, 2=active movement without antigravity, 3=active movement with antigravity, 4=active movement against resistance, and normal=5. The responses for the right side were summed and ranged from 0 to 70.

‡‡Wilson (2012): The *FIM motor score* consists of 13 items that assess function across four different domains of self-care, sphincter control, transfers, and locomotion. The performance level for each item is strictly defined and ranges in value from 1–7, where 1 indicates complete dependence in an activity, and a score of 6 or greater indicates that a patient is capable of performing that activity independently, without supervision or help. The result is a discrete outcome variable with a minimum value of 13 and a maximum value of 91, with a larger value implying superior functional status. *Functional dependence* was a dichotomous variable defined as having achieved/not achieved ≥ 6 for all 13 FIM score items.

**Key Question 2: Association between MRI spinal cord lesion characteristics, length and pattern and neurologic outcomes: results from studies that adjusted for baseline neurologic status**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Author (year)** | **Risk of Bias** | **Follow-up duration** | **Potential prognostic factors** | **Outcome measures\*** | | | | | |
|  | | | | | *Worse neurologic recovery (ASIA motor score)* | *Retaining complete SCI (ASIA impairment scale, grade A)* | *Poor neurologic recovery*†† *(Frankel grade)* | *Worse motor function recovery*‡‡ | *Worse pin prick recovery*‡‡ | *Worse light touch recovery*‡‡ |
| ***Radiographic*** | Aarabi (2011) | Moderately low | Mean: 29.1 ± 27 months | Lower MCC† | *P = .02* |  |  |  |  |  |
| Lower MSCC† | *P = 0.34* |  |  |  |  |  |
| Smaller sagittal diameter of stenotic spinal canal at the point of MSCC (mm) | *P = .02* |  |  |  |  |  |
| Longer parenchymal damage (mm) | *P = 0.25* |  |  |  |  |  |
| ***Clinical*** | Lower admission ASIA motor score | *P = .003* |  |  |  |  |  |
| Mechanism of injury | *P = 0.50* |  |  |  |  |  |
| Number of stenotic skeletal segments | *P = 0.64* |  |  |  |  |  |
| Surgical technique (front, back, circumferential decompression) | *P = 0.09* |  |  |  |  |  |
| Time delay after injury until surgery | *P = 0.478* |  |  |  |  |  |
| ***Demographic*** | Older age | *P =0.08* |  |  |  |  |  |
| ***Radiographic*** | Boldin (2006) | Moderately low | Median: 35 (24-65) months | Presence of cord edema‡ |  | *P = NS* |  |  |  |  |
| Presence of intramedullary hemorrhage‡ |  | *P = NS* |  |  |  |  |
| Longer edema lesion length‡ |  | *P = .02;*  *Each mm increase in edema lesion length increased risk of retaining complete SCI (1.15, 95% CI 1.03 to 1.29); NS when length of hematoma was added to model* |  |  |  |  |
| Longer hemorrhage lesion length‡ |  | *P = .02*§  *Each mm increase in length of hemorrhage lesion increased the risk of retaining complete SCI (1.81, 95% CI (1.09 –3.02)* |  |  |  |  |
| ***Clinical*** | Worse admission ASIA impairment score |  | *P = NS* |  |  |  |  |
| ***Radiographic*** | Miyanji (2007) | Moderately low | Mean: 7.3 (1-35) months | Lower MCC\*\* | *P = NS* |  |  |  |  |  |
| Lower MSCC\*\* | *P = NS* |  |  |  |  |  |
| Presence of cord edema\*\* | *P = NS* |  |  |  |  |  |
| Presence of intramedullary hemorrhage\*\* | *P = 0.002* |  |  |  |  |  |
| Presence of cord swelling\*\* | *P = .054* |  |  |  |  |  |
| Presence of soft-tissue injury\*\* | *P = NS* |  |  |  |  |  |
| Presence of pre-injury stenosis | *P = NS* |  |  |  |  |  |
| Presence of disk herniation | *P = NS* |  |  |  |  |  |
| Longer SCI lesion length\*\* | *P = NS* |  |  |  |  |  |
| ***Clinical*** | Lower baseline ASIA motor score | *P < .001* |  |  |  |  |  |
| Worse baseline ASIA impairment scale score | *P = NS* |  |  |  |  |  |
| ***Radiographic*** | Selden§§ (1999) | Moderately high | Mean: 18.5 (1-74) months | Smaller maximal cross-sectional diameter within swollen length of the cord |  |  | *P = NS* |  |  |  |
| Lower cord compression |  |  | *P = NS* |  |  |  |
| Longer cord swelling length†† |  |  | *P = NS* |  |  |  |
| Longer cord edema length††§§ |  |  | *P =0.071* |  |  |  |
| Intra-axial hematoma§§ |  |  | *P <0.001* |  |  |  |
| Rostrocaudal length of cord hematoma§§ |  |  | *P = 0.028* |  |  |  |
| Cord compression from extra-axial hematoma§§ |  |  | *P = 0.077* |  |  |  |
| Longer cord hematoma length-related signal changes |  |  | *P = NS* |  |  |  |
| ***Clinical*** | Worse admission Frankel grade |  |  | *P < .001* |  |  |  |
| ***Radiographic*** | Shepard (1999) | Moderately low | Mean: 6 weeks | Presence of cord edema‡‡ |  |  |  | *P = .06*  *-3.34 change* | *P = NS*  *-2.46change* | *P = .05*  *-3.41 changge* |
| Presence of hemorrhage‡‡ |  |  |  | *P = NS*  *-2.70 change* | *P = NS*  *-0.63 change* | *P = NS*  ***-0.93change*** |
| Presence of cord contusion§§ |  |  |  | *P = NS*  *-0.36 change* | *P = NS*  *-3.35 change* | *P = NS*  *-1.38 change* |
| ***Clinical*** | Worse baseline pin prick score‡‡ |  |  |  | *P = NS* | *P = NS* | *P = NS* |
| Worse baseline light touch score‡‡ |  |  |  | *P = NS* | *P = NS* | *P = NS* |
| Worse baseline motor function score‡‡ |  |  |  | *P = NS* | *P = NS* | *P = NS* |

ASIA = American Spinal Injury Association; FIM = functional independence measure; MCC = maximum canal compromise; MRI = magnetic resonance imaging; MSCC = maximum spinal cord compression; NS = not significant; SCI = spinal cord injury

\* Results presented as reported by the authors based on multivariate regression analyses, adjusted for baseline neurologic status.

† Aarabi (2011): *MCC* = {1 - Di/ [1/2(Da + Db)]} × 100, where Di is the midsagittal diameter of the spinal canal at the point of maximum compression, Da is the diameter of the spinal canal one segment above the highest level of spinal stenosis, and Db is the diameter of the spinal canal one segment below the lowest level of spinal stenosis. *MSCC* = {1 - di/ [1/2(da + db)]} × 100, where da is the diameter of the spinal cord at a normal segment above the highest level of spinal stenosis, db is the diameter of the spinal cord one segment below the lowest level of stenosis, and di is the diameter of the spinal cord at the level of maximum stenosis.

‡ Boldin (2006): *Cord edema* was defined as lesions associated with cord enlargement; a linear hyperintensity was frequently seen extending superiorly and inferiorly from the lesion. *Intramedullary hemorrhage* was defined as high signal intensity on T1-weighted and low signal intensity on T2-weighted images consistent with hemorrhage, which was centered within a region of cord edema*. Edema lesion length* was measured by the proximal-distal range of the intramedullary hyperintensity on T2-weighted spinal echo images. The *hemorrhage lesion* *length* was measured by the proximal-distal range of the intramedullary signal alterations, meaning high signal on T1-weighted sequences and/or low signal on T2-weighted gradient echo sequences.

Based on how the authors report their results, it appears that multivariate regression was only done in patients with complete SCI via forward selection models including all MRI findings and neurologic scores; among patients with and without hematoma, only length of edema was a significant predictor associated with risk of retaining complete SCI. Based on this information it was inferred that presence of hematoma was not a predictor of the outcome; however, the authors did not provide additional information. When hematoma length was added to the model, it was the only predictive variable in patients with hemorrhage. The subanalysis of ASIA A patients appears to be post-hoc.

§ When length of hematoma was added to the statistical model, the only significant predictor in the model was length of hematoma.

\*\* Miyanji (2007): *MCC/MSCC* was defined as described by Fehlings MG, Rao SC, Tator CH, et al. The optimal radiologic method for assessing spinal canal compromise and cord compression in patients with cervical spinal cord injury. Part II: Results of a multicenter study. Spine 1999;24 (6):605-613. *Intramedullary hemorrhage* was defined as aregion of decreased signal intensity surrounded by a thin rim of high signal intensity on T2-weighted images. *Cord Edema* was defined as aregion of high signal intensity on T2-weighted images. *Cord swelling* was defined as a focal widening of the spinal cord. *Soft-tissue* *injury* was defined as an increased signal intensity of the perivertebral tissues on T2-weighted images. *Spinal cord lesion length* was defined as the distance between the most cephalic and the most caudal extent of the cord signal intensity change on T2-weighted images. Authors used stepwise multivariable regression with the following pre-defined variables to predict ASIA motor score at last follow-up: intramedullary hemorrhage, cord swelling, STI, preinjury stenosis, disc herniation, lesion length MCC and MSCC. In the model adjusted for baseline neurologic assessment, only hemorrhage and cord swelling were retained as significant predictors outcome. Results for the other variables listed were assumed to be not statistically significant.

†† Selden (1999*): Rostrocaudal length of spinal cord swelling* was described as a pathologically increased spinal cord diameter. *Rostrocaudal length of spinal cord edema* wasdescribed as an abnormally increased T2-weighted signal. In univariate correlation analysis, rostrocaudal length of edema was associated with worse Frankel grade; however, the correlation was small (r = -0.28, p =0.036; the association was no longer statistically significant following what is interpreted to be multivariate analysis and adjustment for baseline neurologic status).

‡‡ Shepard (1999): *Spinal cord lesion* patterns were defined as signal intensity or signal changes on T2-weighted images consistent with characteristics of lesions. *Pin prick/light touch score* was determined as follows: 29 spinal cord segments were tested bilaterally for response to pin prick and light touch, each of which were scored at 1=absent, 2=dysfunctional and 3=normal. Sensory scores ranged from a total score of 87 (normal response) to a score of 29 (no response in any segment). Responses for the right side were summed, unless some cord segments could not be tested, in which case responses for the left side were used. *Motor function score* was determined by bilaterally measuring fourteen muscle roots, with 0=no contraction, 1=flicker/trace of contraction, 2=active movement without antigravity, 3=active movement with antigravity, 4=active movement against resistance, and normal=5. The responses for the right side were summed and ranged from 0 to 70.

§§ Selden provides inadequate information regarding multivariate analysis and reports the “independent importance…of MRI findings in predicting ultimate neurological outcomes”, but doesn’t provide models or details and also mentions that results were for the interaction between initial neurologic function and MRI findings; our best estimation is that the results in the “Correlation with neurologic improvement” section represent multivariate analysis.

**Key Question 2: Association between MRI spinal cord lesion characteristics, length and pattern and functional outcomes: results from studies that adjusted for baseline neurologic status**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Author (year)** | **Risk of Bias** | **Follow-up duration** | **Potential prognostic factors** | **Outcome measures\*** | | | | | |
|  | | | | | *Worse FIM score*† | *Functional dependence*† | *Decreased upper extremity motor function*‡ | *Decreased lower extremity motor function*‡ | *Decreased upper extremity muscles with useful function*‡ | *Decreased lower extremity muscles with useful function*‡ |
| ***Radiographic*** | Aarabi (2011) | Moderately low | Mean: 29.1 ± 27 months | Lower MCC§ | *P = .02* |  |  |  |  |  |
| Lower MSCC§ | *P = 0.47* |  |  |  |  |  |
| Lower sagittal diameter of stenotic spinal canal at the point of MSCC (mm) | *P = 0.10* |  |  |  |  |  |
| Longer parenchymal damage (mm) | *P = 0.16* |  |  |  |  |  |
| ***Clinical*** | Lower admission ASIA motor score | *P = .03* |  |  |  |  |  |
| Mechanism of injury | *P = 0.08* |  |  |  |  |  |
| Number of stenotic skeletal segments | *P = 0.24* |  |  |  |  |  |
| Surgical technique (front, back, circumferential decompression) | *P = 0.10* |  |  |  |  |  |
| Time delay after injury until surgery | *P = 0.8* |  |  |  |  |  |
| ***Demographic*** | Older age | *P = .02* |  |  |  |  |  |
| ***Radiographic*** | Wilson (2012) | Moderately low | 6 month: 17.6% (66/376),  12 month: 82.4% (310/376) | MRI signal characteristics consistent with edema or hemorrhage | *P = 0.19* | *P = 0.54*  *OR 0.75(No confidence interval)* |  |  |  |  |
| ***Clinical*** | Initial ASIA motor score ≤ 50† | *P < .01* | *P < .01* |  |  |  |  |
| Lower initial ASIA impairment score | *P < .01* | *P < .01* |  |  |  |  |
| ***Demographic*** | Older age | *P < .01* | *P < .01* |  |  |  |  |
| ***Radiographic*** | Flanders (1996) | Moderately high | Mean: ≥ 12 months | Presence of hemorrhage‡ |  |  | *P < .001* | *P < .001* | *P < .001* | *P < .001* |
| Rostral point of edema‡ |  |  | *P = 0.617* | *P = 0.132* | *P = .011* | *P = 0.126* |
| Longer edema lesion length‡ |  |  | *P = .03* | *P = .03* | *P = 0.85* | *P < .001* |
| ***Clinical*** | Lower initial ASIA motor score |  |  | *P = .002* | *P < .001* | *P < .05* | *P < .001* |

ASIA = American Spinal Injury Association; FIM = functional independence measure; MCC = maximum canal compromise; MRI = magnetic resonance imaging; MSCC = maximum spinal cord compression; NS = not significant;

\* Results presented as reported by the authors based on multivariate regression analyses, adjusted for baseline neurologic status.

† Wilson (2012): *ASIA motor score* was dichotomized at ≤ 50 and > 50. The *FIM motor score* consists of 13 items that assess function across four different domains of self-care, sphincter control, transfers, and locomotion. The performance level for each item is strictly defined and ranges in value from 1–7, where 1 indicates complete dependence in an activity, and a score of 6 or greater indicates that a patient is capable of performing that activity independently, without supervision or help. The result is a discrete outcome variable with a minimum value of 13 and a maximum value of 91, with a larger value implying superior functional status. *Functional dependence* was a dichotomous variable defined as having achieved/not achieved ≥ 6 for all 13 FIM score items.

‡ Flanders (1996): *Presence of hemorrhage* was defined as focal decreased signal on T2-weighted or gradient-echo images. *Edema lesion length* was defined as the length of intramedullary hyperintensity on T2-weighted images. *Rostral point of edema* was determined by locating the longitudinal boundary of spinal cord edema relative to the nearest adjacent spinal vertebral landmark; an anatomic location was labeled as the portion of the closest vertebral body that intersected a horizontal line drawn through the lesion. *Motor function* was assessed by testing key individual muscles in each of the upper and lower extremities. Each muscle received a score of 0 to 5 for a total possible motor index score of 50 for upper extremity function and 40 for lower extremity function. *Minimally useful motor function* in an individual muscle was defined as a score of 3 or better on the 5-point manual muscle test. A score of 3 represents the ability for active movement with a full range of motion against gravity.

§ Aarabi (2011): *MCC* = {1 - Di/ [1/2(Da + Db)]} × 100, where Di is the midsagittal diameter of the spinal canal at the point of maximum compression, Da is the diameter of the spinal canal one segment above the highest level of spinal stenosis, and Db is the diameter of the spinal canal one segment below the lowest level of spinal stenosis. *MSCC* = {1 - di/ [1/2(da + db)]} × 100, where da is the diameter of the spinal cord at a normal segment above the highest level of spinal stenosis, db is the diameter of the spinal cord one segment below the lowest level of stenosis, and di is the diameter of the spinal cord at the level of maximum stenosis.

**Key Question 2: Association between MRI spinal cord lesion characteristics, length and pattern and patient-reported outcomes**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Author (year)** | **Risk of Bias** | **Follow-up duration** | **Potential prognostic factors** | **Outcome measures\*** | |
|  | **Aarabi (2011)** | Moderately low | Mean: 29.1 ± 27 months |  | *Less manual dexterity*† | *Greater dysesthetic pain*† |
| ***Radiographic*** | Lower MCC† | *P = NS* | *P = NS* |
| Lower MSCC† | *P = NS* | *P = NS* |
| Lower sagittal diameter of stenotic spinal canal at the point of MSCC (mm) | *P = NS* | *P = NS* |
| Longer parenchymal damage (mm) | *P = .002* | *P = .04* |
| ***Clinical*** | Lower admission ASIA motor score | *P = .0002* | *P = NS* |
| Mechanism of injury | *P = NS* | *P = NS* |
| Number of stenotic skeletal segments | *P = NS* | *P = NS* |
| Surgical technique (front, back, circumferential decompression) | *P = NS* | *P = NS* |
| Time delay after injury until surgery | *P = NS* | *P = NS* |
| ***Demographic*** | Younger age | *P = NS* | *P = .02* |

MCC = maximum canal compromise; MRI = magnetic resonance imaging; MSCC = maximum spinal cord compromise; NS = not significant

\* Results presented as reported by the authors based on multivariate regression analyses, adjusted for baseline neurologic status.

† Aarabi 2011: *MCC* = {1 - Di/ [1/2(Da + Db)]} × 100, where Di is the midsagittal diameter of the spinal canal at the point of maximum compression, Da is the diameter of the spinal canal one segment above the highest level of spinal stenosis, and Db is the diameter of the spinal canal one segment below the lowest level of spinal stenosis. *MSCC* = {1 - di/ [1/2(da + db)]} × 100, where da is the diameter of the spinal cord at a normal segment above the highest level of spinal stenosis, db is the diameter of the spinal cord one segment below the lowest level of stenosis, and di is the diameter of the spinal cord at the level of maximum stenosis. *Manual dexterity level* was defined as the patient’s own perception of dexterity and skill in managing the following tasks: 1) using a keyboard, 2) playing a musical instrument, 3) buttoning his/her shirt, 4) grooming, and 5) writing. *Dysesthetic pain level* was a subjective rating on a 0 to 10 analog scale.

1*. General Search strategy and history*

A systematic search of PubMed/MEDLINE, EMBASE, the Cochrane Collaboration Library and Google Scholar for literature was conducted with no limits on language. Studies were only considered if they had abstracts and were on humans. The original search was updated twice, most recently to search for studies published through May 12, 2015. In addition to structured database searches, the reference lists of included studies and identified systematic reviews were searched for potentially relevant studies. In addition, studies suggested by clinical experts were evaluated against the inclusion/exclusion criteria.

The general search codes included the following terms/strategy for PubMed/MEDLINE; a similar strategy was used for EMBASE:

General search:

(acute spinal cord injury OR acute spinal cord injuries OR (acute AND spinal cord injuries)) AND magnetic resonance imaging NOT case report,

Expanded search general code:

(Spinal Cord Injuries[MeSH] OR Spinal Cord Compression[MeSH] OR Spinal Cord Ischemia[MeSH] OR Central Cord Syndrome[MeSH] OR (Myelopathy AND (Trauma OR Traumas OR Traumatic OR Post-traumatic OR Posttraumatic)) OR ((Spine OR Spinal) AND (Trauma OR Traumas OR Traumatic OR Injur\* OR Damag\*)) OR (Cord AND (Contusion\* OR Laceration\* OR Transaction\* OR Trauma OR Traumas OR Traumatic\* OR Ischemi\*)) OR Central Cord Injury Syndrome OR Central Spinal Cord Syndrome OR Cervical Vertebrae/injuries[MeSH] OR Lumbar Vertebrae/injuries[MeSH] OR Thoracic Vertebrae/injuries[MeSH] OR Spinal Cord Injuries[MeSH] OR Paraplegia[MeSH] OR Quadriplegia[MeSH] OR Paraplegi\* OR Quadriplegi\* OR Tetraplegi\*) AND acute AND magnetic resonance imaging[MeSH] NOT case report NOT stroke

DTI: Diffusion Tensor Imaging[MeSH] AND spinal cord injuries[MeSH]

(Spinal Cord Injuries[MeSH] OR Spinal Cord Compression[MeSH] OR Spinal Cord Ischemia[MeSH] OR Central Cord Syndrome[MeSH] OR (Myelopathy AND (Trauma OR Traumas OR Traumatic OR Post-traumatic OR Posttraumatic)) OR ((Spine OR Spinal) AND (Trauma OR Traumas OR Traumatic OR Injur\* OR Damag\*)) OR (Cord AND (Contusion\* OR Laceration\* OR Transaction\* OR Trauma OR Traumas OR Traumatic\* OR Ischemi\*)) OR Central Cord Injury Syndrome OR Central Spinal Cord Syndrome OR Cervical Vertebrae/injuries[MeSH] OR Lumbar Vertebrae/injuries[MeSH] OR Thoracic Vertebrae/injuries[MeSH] OR Spinal Cord Injuries[MeSH] OR Paraplegia[MeSH] OR Quadriplegia[MeSH] OR Paraplegi\* OR Quadriplegi\* OR Tetraplegi\*) AND diffusion tensor imaging[MeSH] NOT case report NOT stroke NOT traumatic brain injury NOT review

2. *Data Extraction*

Each retrieved citation was reviewed independently by two reviewers. Most articles were excluded based on information provided in the title or abstract. The full texts were reviewed for citations that appeared to be appropriate or could not be excluded based on the title and abstract. Any disagreement between the two reviewers was resolved through discussion. From the included articles, the following data were extracted: study design; patient demographics; inclusion and exclusion criteria; baseline population/disease characteristics; follow-up duration and the rate of follow-up (if reported or calculable); baseline MRI characteristics and assessment; timing of MRI in relation to injury and intervention (if applicable); treatment groups; reported information on specific treatment decisions; neurological, functional, safety, and quality of life outcomes; prognostic factors evaluated; association between prognostic factors and reported outcomes; and cost-effectiveness data.

3. *Study Quality and Overall Quality (strength) of Evidence*

Each article was evaluated for risk of bias. The method used for assessing quality of individual studies as well as the overall body of evidence incorporates aspects from (1) the rating scheme developed by the Oxford Centre for Evidence-based Medicine[Phillips 2001] with modification by *The Journal of Bone and Joint Surgery American Volume* (*J Bone Joint Surg Am*), [Wright 2003]; (2) precepts outlined by the Grades of Recommendation Assessment, Development and Evaluation (GRADE) Working Group [Atkins 2004; Balshem 2011]; (3) and recommendations made by the Agency for Healthcare Research and Quality (AHRQ) [Methods guide; West 2002]. This appraisal system accounts for features of methodologic quality and important sources of bias by combining epidemiologic principles with characteristics of study design (see Tables 3a-3d below). Qualitative analysis was performed according to HRQ-required and additional domains [OWENS].

*3. Risk of Bias Tables*

*3a. Risk of bias assessment for prognostic studies*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Methodologic principle** | **Aarabi**  **(2011)** | **Boldin**  **(2006)** | **Flanders (1996)** | **Miyanji**  **(2007)** | **Selden**  **(1999)** | **Shepard**  **(1999)** | **Wilson (2012)** |
| Study design |  |  |  |  |  |  |  |
| Prospective cohort study |  | ✓ |  | ✓ |  | ✓ | ✓ |
| Retrospective cohort study | ✓ |  | ✓ |  | ✓ |  |  |
| Case-control study |  |  |  |  |  |  |  |
| Cross-sectional study |  |  |  |  |  |  |  |
| Case-series |  |  |  |  |  |  |  |
| **COHORT STUDIES** |  |  |  |  |  |  |  |
| Patients at similar point in the course of their disease or treatment | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Complete follow-up of > 80% | ✓ |  |  |  |  |  |  |
| Patients followed long enough for outcomes to occur | ✓ | ✓ | ✓ | ✓ | ✓ | ✓‡ | ✓ |
| Accounting for other prognostic factors\* | ✓ | ✓ | ✓ | ✓ | ✓† | ✓ | ✓ |
| **CASE-CONTROL STUDIES** |  |  |  |  |  |  |  |
| Incidence cases from defined population over a specified time period |  |  |  |  |  |  |  |
| Controls represent the population from which the cases come |  |  |  |  |  |  |  |
| Exposure precedes an outcome of interest |  |  |  |  |  |  |  |
| Accounting for other prognostic factors |  |  |  |  |  |  |  |
| **CROSS-SECTIONAL STUDIES** |  |  |  |  |  |  |  |
| A representative sample of the population of interest |  |  |  |  |  |  |  |
| Exposure that precedes an outcome of interest (e.g., sex, genetic factor) |  |  |  |  |  |  |  |
| Accounting for other prognostic factors |  |  |  |  |  |  |  |
| For surveys, a return rate of > 80% |  |  |  |  |  |  |  |
| **Evidence level** | **II** | **II** | **III** | **II** | **III** | **II** | **II** |
| **Risk of bias** | **Moderately low** | **Moderately low** | **Moderately high** | **Moderately low** | **Moderately high** | **Moderately low** | **Moderately low** |

\*Authors must consider other factors that might influence patient outcomes

*Blank cells indicate that the criterion was either not met or that it could not be determined*

*†Authors provided inadequate detail regarding multivariate analysis*

‡Effect sizes provided at 6 weeks; clinicians suggest that this may be adequate time for recovery

*3b. Definition of level of evidence and risk of bias for prognostic studies*

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | **Studies of Prognosis** | |
| **Level** | **Risk of bias** | **Study design** | **Criteria** |
| **I** | **Low risk;**  Study adheres to commonly held tenets of high quality design, execution and avoidance of bias | Good quality cohort\* | * Prospective design * Patients at similar point in the course of their disease or treatment * F/U rate of ≥ 80%† * Patients followed long enough for outcomes to occur * Control for other prognostic factors‡ |
| **II** | **Moderately low risk:**  Study has potential for some bias; does not meet all criteria for level I but deficiencies are not likely to invalidate results or introduce significant bias | Moderate quality cohort | * Prospective design, with violation of one of the criteria for a good quality cohort study * Retrospective design, meeting all the rest of the criteria in level I |
| **III** | **Moderately high risk:**  Study has flaws in design and/or execution that increase potential for bias and may invalidate study results | Poor quality cohort  Good quality case-control or cross-sectional study | * Prospective design with violation of ≥2 criteria for good quality cohort * Retrospective design with violation of ≥1 criteria for good quality cohort * A good case-control study§ * A good cross-sectional study\*\* |
| **IV** | **High risk:**  Study has significant potential for bias; does not include design features geared toward minimizing bias and/or does not have a comparison group | Poor quality case-control or cross-sectional  Case series§ | * Other than a good case-control study * Other than a good cross-sectional study * Any case series†† design |

\*Cohort studies follow individuals with the exposure of interest over time and monitor for occurrence of the outcome of interest.

†Applies to cohort studies only.

‡Authors must consider other factors that might influence patient outcomes and should control for them when appropriate.

§A good case-control study must have the all of the following: all incident cases from the defined population over a specified time period, controls that represent the population where the cases come from, exposure that precedes an outcome of interest, and control for other prognostic factors.

\*\*A good cross-sectional study must have all of the following: a representative sample of the population of interest, an exposure that precedes an outcome of interest (e.g., sex, genetic factor), control for other prognostic factors, and, for surveys, at least an 80% return rate.

††A case-series design for prognosis is one where all the patients in the study have the exposure of interest. Since all the patients have the exposure, risks of an outcome can be calculated only for those with the exposure and cannot be compared to those who do not have the exposure. For example, a case-series evaluating the effect of smoking on spine fusion that only recruits patients who smoke can only provide the risk of pseudoarthorsis in patients who smoke but cannot compare this risk to those that do not smoke.

**Prognostic study list and risk of bias**

* **Aarabi (N=42)** – Moderately low ROB (CoE II); Retrospective
* **Boldin (N = 29)** – Moderately low ROB (CoE II); Prospective, follow-up could not be determined – no information on whether consecutive patients enrolled, number of eligible subjects, number excluded or number followed/lost to follow-up.
* **Flanders (N = 104)** – Moderately high ROB (CoE III); loss to follow-up >80% (79/197 had lost or incomplete images; f/u of 40%); Retrospective
* **Miyanji (N =100)** – Moderately low ROB (CoE II); Prospective, follow-up could not be determined – no information on whether consecutive patients enrolled, number of eligible subjects, number excluded or number followed/lost to follow-up
* **Selden (N = 55)** - Moderately high ROB (CoE III); Retrospective, follow-up could not be determined – no information on whether consecutive patients enrolled, number of eligible subjects, number excluded or number followed/lost to follow-up; authors provided inadequate detail regarding multivariate analysis
* **Shepard (N = 191)** – Moderately low ROB (CoE II); Prospective, follow-up could not be determined (MRI was electively done in 282/499 participants, 191 of whom had MRI within 72 hours)
* **Wilson (N = 376)** - Moderately low ROB (CoE II); Prospective, follow-up could not be determined (data set from two studies, no information on number of eligible subjects, number excluded or number followed/lost to follow-up).

*3c. Level of Evidence and risk of bias criteria for therapeutic studies*

|  |  |
| --- | --- |
| **Methodologic principle** | **Papadopoulos (2002)** |
| **Study design** |  |
| Randomized controlled trial |  |
| Prospective cohort study | ✓ |
| Retrospective cohort study |  |
| Case-control |  |
| Case-series |  |
| Random sequence generation\* |  |
| Statement of concealed allocation\* |  |
| Intention to treat\* |  |
| Independent or blind assessment |  |
| Co-interventions applied equally | N/A |
| Complete follow-up of >80% |  |
| Adequate sample size | ✓ |
| Controlling for possible confounding† | ✓ |
| **Evidence level** | III |
| **Risk of Bias** | Moderately High |

\*Applies only to randomized controlled trials

†Groups must be comparable with respect baseline characteristics or there must be evidence of control for confounding

*Blank cells indicate that the criterion was either not met or that it could not be determined*

*3d. Definition of level of evidence and risk of bias for studies on therapy*

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | **Studies of Therapy** | |
| **Level** | **Bias Risk** | **Study design** | **Criteria** |
| **I** | **Low risk:**  Study adheres to commonly held tenets of high quality design, execution and avoidance of bias | Good quality RCT | * Random sequence generation * Allocation concealment * Intent-to-treat analysis * Blind or independent assessment for important outcomes * Co-interventions applied equally * F/U rate of 80%+ * Adequate sample size |
| **II** | **Moderately low risk:**  Study has potential for some bias; study does not meet all criteria for level I, but deficiencies are not likely to invalidate results or introduce significant bias | Moderate or poor quality RCT | * Violation of one of the criteria for good quality RCT |
|  | Good quality cohort | * Blind or independent assessment in a prospective study, or use of reliable data\* in a retrospective study * Co-interventions applied equally * F/U rate of 80%+ * Adequate sample size * Controlling for possible confounding† |
| **III** | **Moderately High risk:**  Study has significant flaws in design and/or execution that increase potential for bias and may invalidate study results | Moderate or poor quality cohort | * Violation of any of the criteria for good quality cohort |
|  | Case-control | * Any case-control design |
| **IV** | **High risk:**  Study has significant potential for bias; lack of comparison group precludes direct assessment of important outcomes | Case series | * Any case series design |

\* Outcome assessment is independent of healthcare personnel judgment. Reliable data are data such as mortality or re-operation.

† Authors must provide a description of robust baseline characteristics, and control for those that are unequally distributed between treatment groups.

4. Excluded articles.

|  |  |  |
| --- | --- | --- |
| **Author** | **Year** | **Reason for exclusion** |
| *Key Questions 1, 2* | | |
| Aebli, N., Ruegg T.B., et al. (2013). “Predicting the risk and severity of acute spinal cord injury after a minor trauma to the cervical spine.” Spine 13:597-604. | 2013 | Did not assess outcomes of interest |
| Andreoli, C., M. C. Colaiacomo, et al. (2005). "MRI in the acute phase of spinal cord traumatic lesions: Relationship between MRI findings and neurological outcome." Radiol Med 110(5-6): 636-645. | 2005 | Retrospective analysis, no multivariate analysis of the association between risk factors and outcomes |
| Boese C.K., Nerlich M, et al. (2013). “Early magnetic imaging in spinal cord injury without radiological abnormality in adults: A retrospective study.” J Trauma Acute Care Surg 74(3): 845/848. | 2013 | Not population of interest; population includes 26.6% Frankel Grade E |
| Bondurant, F. J., H. B. Cotler, et al. (1990). "Acute spinal cord injury. A study using physical examination and magnetic resonance imaging." Spine (Phila Pa 1976) 15(3): 161-168 | 1990 | No multivariate analysis of the association between risk factors and outcomes |
| Buchberger W, Springer P, Birbamer G, Judmaier W, Kathrein A, Daniaux H. (1995). “[Magnetic resonance tomography in spinal trauma].” Rofo 163(1):53-9. | 1995 | No multivariate analysis of the association between risk factors and outcomes |
| Cotler, H. B., M. V. Kulkarni, et al. (1988). "Magnetic resonance imaging of acute spinal cord trauma: preliminary report." J Orthop Trauma 2(1): 1-4. | 1988 | Population overlap with Bondurant 1990 article; no multivariate analysis of the association between risk factors and outcomes |
| Dai, L. (2001). "Magnetic resonance imaging of acute central cord syndrome: correlation with prognosis." Chin Med Sci J 16(2): 107-110. | 2001 | No multivariate analysis of the association between risk factors and outcomes |
| Davis, S. J. and M. S. Khangure (1994). "A review of magnetic resonance imaging in spinal trauma." Australas Radiol 38(4): 241-253. | 1994 | No multivariate analysis of the association between risk factors and outcomes |
| Endo, T., S. Suzuki, et al. (2011). "Prediction of neurological recovery using apparent diffusion coefficient in cases of incomplete spinal cord injury." Neurosurgery 68(2): 329-336. | 2011 | Did not assess prognostic factors of interest |
| Flanders, A. E., C. M. Spettell, et al. (1999). "The relationship between the functional abilities of patients with cervical spinal cord injury and the severity of damage revealed by MR imaging." AJNR Am J Neuroradiol 20(5): 926-934. | 1999 | No multivariate analysis of the association between risk factors and outcomes |
| Freund, P., N. Weiskopf, et al. (2013). "MRI investigation of the sensorimotor cortex and the corticospinal tract after acute spinal cord injury: A prospective longitudinal study." The Lancet Neurology 12(9): 873-881. | 2013 | Did not assess prognostic factors of interest |
| Goradia, D., K. F. Linnau, et al. (2007). "Correlation of MR imaging findings with intraoperative findings after cervical spine trauma." AJNR Am J Neuroradiol 28(2): 209-215. | 2007 | Did not assess outcomes of interest |
| Huang, Y. H., T. M. Yang, et al. (2009). "The prognosis of acute blunt cervical spinal cord injury." J Trauma 66(5): 1441-1445. | 2009 | Did not assess prognostic factors of interest |
| Huang YH, Ou CY. Magnetic resonance imaging predictors for respiratory failure after cervical spinal cord injury. *Clinical Neurology and Neurosurgery.* 2014;126:30-34. | 2014 | Did not assess outcomes of interest |
| Keiper, M. D., R. A. Zimmerman, et al. (1998). "MRI in the assessment of the supportive soft tissues of the cervical spine in acute trauma in children." Neuroradiology 40(6): 359-363. | 1998 | Did not assess outcomes of interest for KQ1; did not meet inclusion criteria for KQ3 |
| Kulkarni, M. V., C. B. McArdle, et al. (1987). "Acute spinal cord injury: MR imaging at 1.5 T." Radiology 164(3): 837-843. | 1987 | Unclear reporting of outcomes, no multivariate analysis of the association between risk factors and outcomes |
| Kulkarni, M. V., F. J. Bondurant, et al. (1988). "1.5 tesla magnetic resonance imaging of acute spinal trauma." Radiographics 8(6): 1059-1082. | 1988 | Unclear reporting of outcomes, no multivariate analysis of the association between risk factors and outcomes |
| Kumar, R., S. Arora, et al. (2011). "Cervical Spine Injury Recovery Prediction Scale: a means of predicting neurological recovery in patients with acute subaxial cervical spine injury." J Orthop Surg (Hong Kong) 19(1): 25-29. | 2011 | Did not assess for baseline neurologic function in multivariate analysis; did not directly address KQs; unclear reporting of outcome variable. |
| Labattaglia, M. P., P. A. Cameron, et al. (2007). "Clinical outcomes of magnetic resonance imaging in blunt cervical trauma." Emerg Med Australas 19(3): 253-261. | 2007 | Did not meet inclusion criteria for KQ3 |
| 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 22:611–621, 2015 | 2015 | Did not assess outcomes of interest or link expansion longitudinally with outcomes of interest |
| Leypold, B. G., A. E. Flanders, et al. (2008). "The early evolution of spinal cord lesions on MR imaging following traumatic spinal cord injury." AJNR Am J Neuroradiol 29(5): 1012-1016. | 2008 | Excluded at title/abstract; did not address KQ or assess outcome of interest; no assessment beyond initial evaluation |
| Liao, C. C., T. N. Lui, et al. (2005). "Spinal cord injury without radiological abnormality in preschool-aged children: correlation of magnetic resonance imaging findings with neurological outcomes." J Neurosurg 103(1 Suppl): 17-23. | 2005 | Did not meet inclusion criteria (N<10) |
| Mahmood, N. S., R. Kadavigere, et al. (2008). "Magnetic resonance imaging in acute cervical spinal cord injury: a correlative study on spinal cord changes and 1 month motor recovery." Spinal Cord 46(12): 791-797. | 2008 | Did not meet inclusion criteria (utilized low-resolution MRI) |
| Marciello, M. A., A. E. Flanders, et al. (1993). "Magnetic resonance imaging related to neurologic outcome in cervical spinal cord injury." Arch Phys Med Rehabil 74(9): 940-946. | 1993 | No multivariate analysis of the association between risk factors and outcomes |
| Miranda, P., P. Gomez, et al. (2008). "Acute traumatic central cord syndrome: analysis of clinical and radiological correlations." J Neurosurg Sci 52(4): 107-112; discussion 112. | 2008 | No multivariate analysis of the association between risk factors and outcomes |
| O'Beirne, J., N. Cassidy, et al. (1993). "Role of magnetic resonance imaging in the assessment of spinal injuries." Injury 24(3): 149-154 | 1993 | No multivariate analysis of the association between risk factors and outcomes |
| Parashari, U. C., S. Khanduri, et al. (2011). "Diagnostic and prognostic role of MRI in spinal trauma, its comparison and correlation with clinical profile and neurological outcome, according to ASIA impairment scale." J Craniovertebr Junction Spine 2(1): 17-26. | 2011 | Multivariate analysis did not control for baseline neurologic function |
| Patel, A. A., R. J. Hurlbert, et al. (2010). "Classification and surgical decision making in acute subaxial cervical spine trauma." Spine (Phila Pa 1976) 35(21 Suppl): S228-234. | 2010 | Did not meet inclusion criteria for KQ3 |
| Poonnoose, P. M., G. Ravichandran, et al. (2002). "Missed and mismanaged injuries of the spinal cord." Journal of Trauma - Injury, Infection and Critical Care 53(2): 314-320. | 2002 | Did not assess comparators of interest for KQ3 |
| Sato, T., S. Kokubun, et al. (1994). "Prognosis of cervical spinal cord injury in correlation with magnetic resonance imaging." Paraplegia 32(2): 81-85. | 1994 | No multivariate analysis of the association between risk factors and outcomes |
| Schaefer, D. M., A. E. Flanders, et al. (1992). "Prognostic significance of magnetic resonance imaging in the acute phase of cervical spine injury." J Neurosurg 76(2): 218-223. | 1992 | No multivariate analysis of the association between risk factors and outcomes |
| Shimada, K. and T. Tokioka (1999). "Sequential MR studies of cervical cord injury: correlation with neurological damage and clinical outcome." Spinal Cord 37(6): 410-415. | 1999 | No multivariate analysis of the association between risk factors and outcomes |
| Shin, J. C., D. Y. Kim, et al. (2005). "Neurologic recovery according to early magnetic resonance imaging findings in traumatic cervical spinal cord injuries." Yonsei Med J 46(3): 379-387. | 2005 | No multivariate analysis of the association between risk factors and outcomes |
| Silberstein, M., D. Brown, et al. (1992). "Suggested MRI criteria for surgical decompression in acute spinal cord injury. Preliminary observations." Paraplegia 30(10): 704-710 | 1992 | Did not meet inclusion criteria (utilized low-resolution MRI) |
| Silberstein, M., B. M. Tress, et al. (1992). "Prediction of neurologic outcome in acute spinal cord injury: the role of CT and MR." AJNR Am J Neuroradiol 13(6): 1597-1608. | 1992 | Did not meet inclusion criteria (utilized low-resolution MRI) |
| Song, J., J. Mizuno, et al. (2006). "Clinical evaluation of traumatic central cord syndrome: emphasis on clinical significance of prevertebral hyperintensity, cord compression, and intramedullary high-signal intensity on magnetic resonance imaging." Surg Neurol 65(2): 117-123. | 2006 | No multivariate analysis of the association between risk factors and outcomes |
| Tewari, M. K., D. S. Gifti, et al. (2005). "Diagnosis and prognostication of adult spinal cord injury without radiographic abnormality using magnetic resonance imaging: analysis of 40 patients." Surg Neurol 63(3): 204-209; discussion 209 | 2005 | No multivariate analysis of the association between risk factors and outcomes |
| Thangarajah, T., D. O'Donoghue, et al. (2011). "Today or tomorrow? A retrospective analysis of the clinical indications used to request urgent magnetic resonance imaging of the spine." Ann R Coll Surg Engl 93(1): 76-80. | 2011 | Population does not meet inclusion criteria |
| Tsuchiya, K., A. Fujikawa, et al. (2006). "Value of diffusion-weighted MR imaging in acute cervical cord injury as a predictor of outcome." Neuroradiology 48(11): 803-808. | 2006 | No multivariate analysis of the association between risk factors and outcomes |
| Wilson, J. R., P. M. Arnold, et al. (2012). "Clinical prediction model for acute inpatient complications after traumatic cervical spinal cord injury: a subanalysis from the Surgical Timing in Acute Spinal Cord Injury Study." J Neurosurg Spine 17(1 Suppl): 46-51. | 2012 | Did not assess predictors or comparators of interest |
| Zhu L, Wu G, Zhou X, Li J, Wen Z, Lin F. Altered spontaneous brain activity in patients with acute spinal cord injury revealed by resting-state functional MRI. *PLoS ONE.* 2015;10(3). | 2015 | Did not assess outcomes of interest |
| *Key Question 3* | | |
| Cheran, S., K. Shanmuganathan, et al. (2011). "Correlation of MR diffusion tensor imaging parameters with ASIA motor scores in hemorrhagic and nonhemorrhagic acute spinal cord injury." J Neurotrauma 28(9): 1881-1892 | 2011 | No analysis of the association between prognostic factors of interest and outcomes of interest. |
| Gustin, S. M., P. J. Wrigley, et al. (2010). "Brain anatomy changes associated with persistent neuropathic pain following spinal cord injury." Cereb Cortex 20(6): 1409-19. | 2010 | Population did not meet inclusion criteria |
| Kim SY, Shin MJ, Chang JH, et al. Correlation of diffusion tensor imaging and phase-contrast MR with clinical parameters of cervical spinal cord injuries. *Spinal Cord.* 2015. | 2015 | Did not assess outcomes of interest |
| Shanmuganathan, K., R. P. Gullapalli, et al. (2008). "Diffusion tensor MR imaging in cervical spine trauma." AJNR Am J Neuroradiol 29(4): 655-659. | 2008 | Did not assess outcomes of interest |
| Sung JK, Jee WH, Jung JY, et al. Differentiation of acute osteoporotic and malignant compression fractures of the spine: Use of additive qualitative and quantitative axial diffusion-weighted MR imaging to conventional MR imaging at 3.0 T. *Radiology.* 2014;271(2):488-498. | 2014 | Did not assess outcomes of interest |
| Vedantam, A., G. Eckardt, et al. (2013). "Clinical Correlates of High Cervical Fractional Anisotropy in Acute Cervical Spinal Cord Injury." World Neurosurg. | 2013 | Did not assess outcomes of interest |
| Vedantam A, Eckardt G, Wang MC, Schmit BD, Kurpad SN. Clinical Correlates of High Cervical Fractional Anisotropy in Acute Cervical Spinal Cord Injury. *World Neurosurgery.* 2015. | 2015 | Did not assess outcomes of interest |
| *Key Question 4* | | |
| Resnick S, Inaba K, Karamanos E, et al. Clinical relevance of magnetic resonance imaging in cervical spine clearance: A prospective study. *JAMA Surgery.* 2014;149(9):934-939. | 2015 | Not a comparative study |
| Selden N, Patel N, Grube S, Papadopoulos S. Emergent MRI and Surgical Decompression Improves Outcome and Reduces Cost in Spinal Cord Injury. Neurosurgery 1997;41(3):725-726. | 1997 | Meeting abstract; not a full economic study |

*5. Appraisal of Bozzo Systematic review.*

**The Role of Magnetic Resonance Imaging in the Management of Acute Spinal Cord Injury**

Anthony Bozzo, Judith Marcoux, Mohan Radhakrishna, Julie Pelletier, and Benoit Goulet

Journal of Neurotrauma 2011 28 8, 1401 -1411

Two systematic review methodologists independently rated this review using the AMSTAR evaluation tool (below). It received 5/11 points for methodologic quality (poor to moderate quality) and was considered to have moderate to moderately high risk of bias.

Appraisal notes:

* Determination of the “overall quality of evidence” for their recommendations is not articulated; based on the types of studies included and the mean Down/Black scores provided (highest was 22/44 points), the overall quality of included studies appears to be poor (case series).
* The basis for determining strength of recommendation is not articulated
* No comprehensive listing of studies (included or excluded) is provided. It is concerning that some tables report “select” papers – it is unclear to what extent this may influence reported findings or recommendations
* No description of study populations, patient characteristics, injury severity or similar factors is provided
* No details of data abstraction or evidence synthesis are provided
* No description of how MRI impacted clinical decision making is provided
* No description of how MRI predicted clinical outcomes is provided; limited graphic information for changes between initial and final ASIA score is provided but no statistical analysis of changes adjusted for baseline neurologic status is provided. Furthermore, timing of follow-up is not described.
* It is not clear from the results presented that the review’s objectives/clinical questions were addressed by the summarized evidence.
* If the GRADE approach was applied to the outcomes described, the overall quality of evidence would likely be very low. This would be an important consideration if this review is used to formulate recommendations

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **AMSTAR Checklist (please assign each criteria 1 point)** | | **Yes** | **No** | **N/A** | **Notes** |
| **1** | **Was an 'a priori' design provided?** | 1 |  |  |  |
| *The research question and inclusion criteria should be established before the conduct of the review.* |
| **2** | **Was there duplicate study selection and data extraction?** *Were there at least two independent data extractors and a consensus procedure for disagreements?* |  | 1 |  | Consensus procedure for disagreements not stated (no/unclear); 2 reviewers independently assessed studies for inclusion (yes) |
| **3** | **Was a comprehensive literature search performed?** *Were at least two electronic sources searched? Were search dates included? Were key words/MeSH terms stated?* | 1 |  |  |  |
| **4** | **Was the status of publication (i.e. grey literature) used as an inclusion criterion?** *Did authors state they searched for reports regardless of publication type? Did they explicitly state what reports were excluded based on publication status, language, etc?* |  | 1 |  | No statement related to publication type; no explicit statement regarding exclusion based on publication status - however, search was limited to articles published in English, and authors stated that case-reports and non-original papers (i.e., reviews) were excluded. References from original articles were searched. |
| **5** | **Was a list of studies (included and excluded) provided?** *Is a complete list of both included/excluded studies provided (likely as a table)?* |  | 1 |  | Incomplete citing of included studies (118 total); no complete list provided or able to be determined; no list of excluded studies provided. Table 2 describes "Select" papers were included with no rationale for how they were selected. |
| **6** | **Were the characteristics of the included studies provided?** *Does an aggregated form such as a table exist to provide data from original studies based on participants, interventions and outcomes? Some range of characteristics included might be age, race, sex, etc.* | 1 |  |  | Very limited information is provided; no information was reported on patient populations, characteristics (aside from age) or various clinical factors |
| **7** | **Was the scientific quality of the included studies assessed and documented?** | 1 |  |  | Level of evidence assessed according to Straus et al 2005 (all articles rated LoE IV, case series); methodological quality assessed using the Downs and Black scoring system; credit given; Mean ratings for D/B by topic ranged from 16.5/44 possible points to 22.7/44 possible points |
| *Did authors state take into account study type and design (e.g. randomization, blinding, placebo controlled studies)?* |
| **8** | **Was the scientific quality of the included studies used appropriately in formulating conclusions?** *Did authors discuss the scientific quality of their included studies in their conclusions?* |  | 1 |  | See the first paragraph under "Discussion and Recommendations." Insufficient details were provided regarding study limitations (Downs and Black system) or how they translated to "weak" or "moderate" evidence in the instances discussed in the conclusions (esp. given that all studies were case series). |
| **9** | **Were the methods used to combine the findings of studies appropriate?** *Did authors assess homogeneity (i.e. Chi-squared test), and take appropriate meta-analyses based on such?* |  |  | X | No meta-analysis; clinically heterogeneous studies; methods of synthesis are not well described. |
|  |
| **10** | **Was the likelihood of publication bias assessed?** *Did authors include graphical (e.g. funnel plot) and/or statistical tests (e.g. Egger regression test) to evaluate included study biases?* |  | 1 |  | No description of consideration of publication bias |
| **11** | **Was the conflict of interest stated?** *Did authors explicitly state any possible conflicting sources of support from included studies?* | 1 |  |  |  |
|  | **Total Score** | **5** | **out of** | **11** |  |