

Detecting Facet Joint and Lateral Mass Injuries of the Subaxial Cervical Spine in Major Trauma Patients

Joost Johannes van Middendorp^{1,2}, Ian Cheung³, Kristian Dalzell⁴, Hamish Deverall⁵, Brian J.C. Freeman^{6,7}, Stephen A.C. Morris⁸, Simon J.I. Sandler⁶, Richard Williams^{9,10}, Y.H. Yau⁶, Ben Goss¹¹

¹Stoke Mandeville Spinal Foundation, National Spinal Injuries Centre, Stoke Mandeville Hospital, Aylesbury, UK

²Harris Manchester College, University of Oxford, Oxford, UK

³Logan Hospital, Logan City, QLD, Australia

⁴Department of Orthopaedics, Christchurch Hospital, Christchurch, New Zealand

⁵Department of Orthopaedics, Waikato Hospital, Hamilton, New Zealand

⁶The Spinal Unit, Department of Orthopaedics and Trauma, Royal Adelaide Hospital, Adelaide, SA, Australia

⁷The University of Adelaide, Adelaide, SA, Australia

⁸Department of Orthopaedics Great Ormond Street Hospital, London, UK

⁹Institute of Health and Biomedical Innovation, Queensland University of Technology, Kelvin Grove, QLD, Australia

¹⁰Trauma Service, Princess Alexandra Hospital, Brisbane, QLD, Australia

¹¹AOSpine Reference Centre, Institute of Health and Biomedical Innovation, Queensland University of Technology, Kelvin Grove, QLD, Australia

Study Design: Radiologic imaging measurement study.

Purpose: To assess the accuracy of detecting lateral mass and facet joint injuries of the subaxial cervical spine on plain radiographs using computed tomography (CT) scan images as a reference standard; and the integrity of morphological landmarks of the lateral mass and facet joints of the subaxial cervical spine.

Overview of Literature: Injuries of lateral mass and facet joints potentially lead to an unstable subaxial cervical spine and concomitant neurological sequelae. However, no study has evaluated the accuracy of detecting specific facet joint injuries.

Methods: Eight spinal surgeons scored four sets of the same, randomly re-ordered, 30 cases with and without facet joint injuries of the subaxial cervical spine. Two surveys included conventional plain radiographs series (test) and another two surveys included CT scan images (reference). Facet joint injury characteristics were assessed for accuracy and reliability. Raw agreement, Fleiss kappa, Cohen's kappa and intraclass correlation coefficient statistics were used for reliability analysis. Majority rules were used for accuracy analysis.

Results: Of the 21 facet joint injuries discerned on CT scan images, 10 were detected in both plain radiograph surveys (sensitivity, 0.48; 95% confidence interval [CI], 0.26–0.70). There were no false positive facet joint injuries in either of the first two X-ray surveys (specificity, 1.0; 95% CI, 0.63–1.0). Five of the 11 cases with missed injuries had an injury below the lowest visible articulating level on radiographs. CT scan images resulted in superior inter- and intra-rater agreement values for assessing morphologic injury characteristics of facet joint injuries.

Conclusions: Plain radiographs are not accurate, nor reliable for the assessment of facet joint injuries of the subaxial cervical spine. CT scans offer reliable diagnostic information required for the detection and treatment planning of facet joint injuries.

Keywords: Zygapophyseal joint; Spine; Spinal injuries; Diagnostic imaging; Reproducibility of results

Received Nov 4, 2014; Revised Nov 4, 2014; Accepted Nov 25, 2014

Corresponding author: Joost Johannes van Middendorp

Stoke Mandeville Spinal Foundation, National Spinal Injuries Centre, Stoke Mandeville Hospital, Aylesbury, Buckinghamshire, HP21 8AL, UK

Tel: +44-1296316783, Fax: +44-1296316783, E-mail: jvanmiddendorp@gmail.com

Introduction

Although the proportion of cervical spine injuries in the adult major trauma population is relatively small (3.5%) [1], missed diagnosis of such an injury could result in serious disability [2]. Routine computed tomography (CT) scans of the spine in the diagnostic work-up of trauma patients is implemented in approximately 80% of level I trauma centers in order to reduce the number of missed injuries [3,4]. However, a recent report demonstrated that clinicians continue to use radiographs of the cervical spine in low-risk adult trauma patients [5].

The lateral masses, superior and inferior articular processes of the subaxial cervical spine form the central pillars of anteroposterior (transverse) and rotational stability [6]. Injuries of these structures potentially lead to an unstable subaxial cervical spine and concomitant neurological sequelae. While several reports on the accuracy of detecting general cervical spine injuries on plain radiographs and CT scans have been published [7-9], no study has evaluated the accuracy of detecting specific facet joint injuries.

Moreover, no spinal injury classification including specific morphological characteristics of the cervical spinal facet joints and lateral masses has yet been accepted and endorsed worldwide. In 2007, Dvorak et al. [10] grouped 90 unilateral facet joint injuries into 9 morphological categories. Subsequently in 2009, Lee and Sung [6] proposed a categorization of lateral mass and facet joint fractures into 6 groups. Both studies, however, did not include bilateral facet joint injuries nor did they present agreement data.

The aim of the current study was twofold. First, we assessed the accuracy of detecting facet joint and lateral mass injuries of the subaxial cervical spine on plain radiographs using CT as a reference standard. Secondly, we assessed the agreement on judging morphological landmarks of the facet joints and lateral masses of the subaxial cervical spine. We conducted a radiologic imaging measurement study for this purpose.

Materials and Methods

1. Setting

Eight spinal surgeons attended a 1-day meeting to score 4 sets of randomly ordered cases with and without facet joint injuries of the cervical spine. The study was divided

into 4 surveys: surveys 1 and 2 were a test and re-test of the same set of 30 randomly (re-)ordered series of conventional (i.e., anteroposterior and lateral) radiographic measurements; surveys 3 and 4 were a test and re-test of the same set of 30 randomly (re-)ordered series of CT scan images. Before the start of survey 1, a study outline describing the rationale and the definitions of the measurements to be made throughout the day was presented. Also, 2 training cases were completed under the guidance of the senior authors (JJvM, BG).

2. Sample definition: patients

Cases were selected from a database comprising a consecutive series of high-energy blunt trauma patients admitted to a Dutch level 1 trauma centre serving a population of over 2.0 million inhabitants [11]. In this study, all admitted high-energy blunt trauma patients older than 15 years were included and underwent routine CT imaging of the cervical spine, thorax and abdomen after clinical evaluation and conventional radiographic work-up.

Included patients were those suffering from life-threatening vital conditions, showing signs of severe injuries during physical examination and/or involved in high-energy injury mechanisms as presented in Table 1. Patients were excluded from the protocol in case of (1) class 3 or 4 shock requiring immediate surgical intervention; (2) suspected or known pregnancy; and (3) neurological condition or deterioration requiring immediate cerebral CT evaluation without any diagnostic delay.

Routine CT protocol was conducted with a Somatom Sensation 16-slice multi-detector CT scanner with automated tube current modulation (Siemens Medical Systems, Erlangen, Germany). CT scans were executed at a tube potential of 120 kV, with a reference value of effective tube current time product of 200 mAs. The detector configuration was 16×1.5 mm. Reconstructed section thickness was 3 mm for bone reconstruction kernel, with an increment of 1.5 mm. Sagittal and coronal reformatted images of the spine were also presented to the raters.

Radiological registrars evaluated all radiological examinations, documented, and entered them into the database under supervision of a trauma-dedicated radiologist. Patients with a subaxial cervical spine injury were retrieved from the database by 1 of the authors (JJvM). Along with the radiology reports, all cases were re-evaluated and categorized as having 'clinically insignificant' or 'clinically

Table 1. Inclusion criteria for the high-energy, blunt trauma study population from which a study sample was drawn

Patients fulfilling one or more of the following criteria	Definitions
Vital conditions <ul style="list-style-type: none"> - Airway patency problems - Breathing problems - Circulatory problems - Neurological problems 	<ul style="list-style-type: none"> - As assessed by anesthesiologist - Breathing frequency: ≥ 30/min - Heart rate ≥ 120/min - Systolic blood pressure < 100 mm Hg, Refill > 4 sec - Exterior blood loss > 500 mL - Glasgow Coma Scale ≤ 13
Physical examination <ul style="list-style-type: none"> - Clinically evident fractures of ≥ 2 long bones - Clinically evident pelvic ring fracture - Signs of unstable vertebral fractures or neural cord compression - Signs of a flail chest/multiple rib fractures 	<ul style="list-style-type: none"> - As assessed by treating physician
Mechanism of injury <ul style="list-style-type: none"> - High-energy mechanism of injury as determined by pre-hospital emergency medical services 	<ul style="list-style-type: none"> - Fall from height > 3 m - Motor vehicle accident ≥ 50 km/hr - Ejection from vehicle - Car rollover - Cabin shortening ≥ 50 cm - Struck pedestrian ≥ 10 km/hr - Struck (motor)cyclist ≥ 30 km/hr - Compressed underneath heavy object

significant' injuries. Subaxial cervical spine injuries were considered 'clinically insignificant' if the injuries were isolated and involved ≥ 1 transverse process(es) (without involvement of the facet joint), spinous process(es) (without involvement of the lamina) or osteophytes. After exclusion of the cases with 'clinically insignificant' injuries, all remaining cases were arbitrarily assigned numbers. Subsequently, 30 cases were randomly selected using an online random integer sequence generator (www.random.org/integers). As a preferred ratio of 2 out of 3 cases with a facet joint injury was not met, 5 random cases without facet joint injury were replaced by 5 manually selected cases with facet joint injury. This resulted in a final cohort of 30 cases with significant subaxial cervical spine injury, of which 21 cases had a facet joint injury and 9 cases had none. The final cohort was arbitrarily re-assigned new numbers and an online random integer set generator (www.random.org/integer-set) was used to generate 4 sets of randomly ordered cases: 2 for the conventional radiography examination (survey 1 & 2) and 2 for the CT evaluation (survey 3 & 4).

3. Sample definition: raters

Eight raters participated in the study. Four raters were

fellowship-trained spine surgeons involved in the management of cervical spine injuries and the remaining 4 raters were orthopedic surgeons who were participating in a spinal surgery fellowship at the time of the study. Each rater was presented the same 4 sets of the randomly ordered cases. The raters were blinded to all other clinical information and none of the raters had any previous clinical information on the cases presented. All cases were scored individually and independently from other raters.

4. Measurements

De-identified images were analyzed using ClearCanvas Workstation V2.0 (Toronto, Canada). The following measurements were made:

1) Visibility of most caudal level (survey 1 & 2)

Before discerning the presence of an injury, raters were asked to report the most caudal spinal level visible on the lateral radiograph. Each level was documented as an articulating level, ranging from C2–C3 to C7–T1 (C, cervical; T, thoracic).

2) Facet joint injury

Raters were asked whether a facet joint injury of the

Table 2. A morphological categorization of lateral mass and facet joint injuries applied in this study

Type	Description
1	Unilateral injury
1.1	Unilateral fracture, without dislocation
1.2	Unilateral dislocation, without fracture
1.3	Unilateral dislocation, with fracture
2	Bilateral injury
2.1	Bilateral fracture, without dislocation
2.2	Bilateral dislocation, without fracture
2.3	Bilateral dislocation, with fracture
3	Lateral mass injury
3.1	Unilateral lateral mass dislocation (floating lateral mass: including pedicle fracture)
3.2	Bilateral lateral mass dislocation (floating lateral mass: including pedicle fracture)

subaxial spine was present or absent. If an injury was detected, the following injury characteristics were measured and reported.

3) Level of injury

Each level was documented as an articulating level, ranging from C2–C3 to C7–T1. In cases where ≥ 1 articulating level was affected, the rater was asked to report both the most severely affected level, as well as the other levels of injury.

4) Injury characteristics

Since there is no widely accepted, endorsed classification of subaxial cervical spinal injuries including specific morphological characteristics of the facet joints and lateral masses, raters categorized each articulating level of injury according to the injury characteristics listed in Table 2. This morphological categorization was created in line with the spinal injury classification principles outlined in a recent paper [12]. Only 1 injury characteristic could be assigned to each level. When ≥ 1 characteristic was possible, the most severe (i.e., highest number) was documented.

5) Vertebral body translation

Vertebral body translation was measured for each articulating level of injury. The measurement technique reported by Bono et al. [13] was used for this purpose (Fig. 1). From a lateral perspective, a posterior tangent line is drawn along the injured and caudal vertebral bodies. The distance between these lines is measured at the level of the inferior aspect of the upper vertebra. Bono et al. [13] reported that when this measurement is made on sagit-

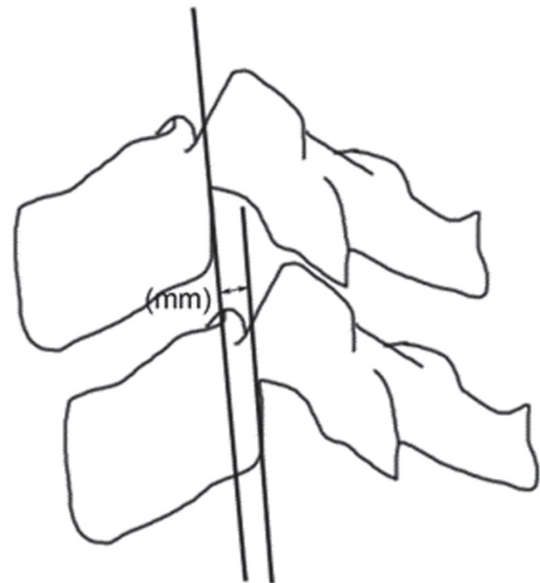


Fig. 1. Measurement technique for vertebral body translation. From a lateral perspective, a posterior tangent line is drawn along the vertebral bodies. The distance between these lines is measured at the level of the inferior aspect of the upper vertebra.

tal CT images, the degree of translation might vary from the left to the right side. Therefore, it is recommended to record the maximal measured translation, since this best reflects the severity of the injury.

5. Statistics

Accuracy is the degree to which the measurement actually

Table 3. Number and proportion of cases with visible articulating levels of the subaxial cervical spine on plain lateral X-rays as assessed in the first two surveys using median values (30 cases, 8 raters)

Articulating level: subaxial cervical spine	Visibility, n (%)	
	Survey 1	Survey 2
C2–C3	30 (100)	30 (100)
C3–C4	30 (100)	30 (100)
C4–C5	29 (97)	30 (100)
C5–C6	28 (93)	28 (93)
C6–C7	23 (77)	23 (77)
C7–T1	11 (37)	12 (40)

C, cervical spine; T, thoracic.

represents what is intended. The accuracy of a test is best assessed by comparison to a reference standard technique that is accurately representative, when possible. We considered conventional radiographs as the ‘test’ and CT images as the reference standard. However, since a gold reference standard was missing, the most likely distribution of true imaging outcomes in the sample was estimated using different majority rules (>50%, >75%, and 100%).

Precision, or reliability, is the extent to which repeated measurements of the same case under similar conditions agree with each other. Observer variation is the most frequently assessed type of reliability and can be divided into 2 components: inter-rater reliability and intra-rater reliability. Inter-rater reliability assesses the reliability, or agreement, of identifying a facet joint or lateral mass injury when measured by different people under similar conditions. Intra-rater reliability assesses the reliability, or reproducibility, of injury identification when measured more than once by the same rater.

The inter-rater reliability of nominal and dichotomous variables was estimated using Fleiss Kappa [14]. The Fleiss Kappa is a multi-rater reliability statistic that is analogous to Cohen’s Kappa that is applicable to 2 raters only. The intra-rater reliability for nominal and dichotomous variables was estimated as a Cohen’s Kappa for the test/re-test pair for each observer. Kappa values range from –1.0 to 1.0, with negative values indicating disagreement and a positive value of 1.0 representing perfect agreement. A 0 value represents no agreement beyond what would be expected by chance. Fleiss [15] proposed the following categories for strength of agreement for the kappa coefficient: ≤ 0.4 =poor, 0.4 – 0.75 =fair to good, and ≥ 0.75 – 1 =excellent agreement. Although this categorization is commonly

applied, no consensus exists as to what constitutes an acceptable kappa coefficient [16]. The intraclass correlation coefficient was used to estimate correlation between all raters for the scale data. Pair-wise agreement coefficients were pooled using random effects model with a 95% confidence interval (CI).

Injury specific measurements could only be calculated for those cases that had 100% agreement on the presence of a facet joint injury. All statistical analyses were considered significant when alpha was <0.05 . Data analyses were performed using the SPSS ver. 16.0 (SPSS, Chicago, IL, USA) and MetaXL ver. 1.3 available from www.epigear.com.

Results

1. Patient characteristics

Thirty cases were scored of which 21 had a facet joint injury and 9 did not. There were 21 male patients and 9 female patients. The mean patient age at the time of injury was 47 years (range, 17 to 93 years).

2. Visibility of most caudal level (survey 1 & 2)

The visibility of each level of the subaxial cervical spine on lateral plain X-rays was presented in Table 3. The articulating level between C3 and C4 was visible in all cases, whereas the cervicothoracic junction was visible in 11 (37%) and 12 (40%) of the 30 cases in surveys 1&2, respectively. The raw agreement on the lowest visible level of the cervical spine was 70% ($\kappa 0.54$; 95% CI, 0.47–0.60) and 71% ($\kappa 0.56$; 95% CI, 0.49–0.62) for surveys 1 and 2, respectively.

Table 4. Number and proportion of cases reaching a threshold of raw agreement on the presence or absence of a facet joint injury of the subaxial cervical spine in the 4 surveys (30 cases, 8 raters) using three different majority rules: >50%, >75%, and 100%

Survey	Raw rater agreement, n (%)		
	>50%	>75%	100%
1. Plain film (test)	25 (83)	22 (73)	9 (30)
2. Plain film (re-test)	28 (93)	24 (80)	13 (43)
3. Computed tomography (test)	29 (97)	29 (97)	20 (67)
4. Computed tomography (re-test)	30 (100)	29 (97)	20 (67)

3. Presence of facet joint injury: accuracy

In the absence of a gold reference standard, the most likely outcomes were estimated with 3 different majority rules (Table 4). Except for 1 case, there was 75% agreement on the presence or absence of a facet joint injury in all cases using CT scan images. Hence, the 75% agreement value was used as a proxy reference standard to assess the accuracy of facet joint injury detection on plain radiographs.

Of the 21 facet joint injuries discerned on CT images, 10 were detected in both plain X-ray surveys (sensitivity, 0.48; 95% CI, 0.26–0.70) by the majority of the raters. There were no false positive facet joint injuries in either of the first 2 X-ray surveys (specificity, 1.0; 95% CI, 0.63–1.0). Eleven of the 21 cases (52%) with facet joint injuries seen on CT scan images were missed in the plain radiograph series. Five cases had an injury below the lowest visible articulating level, on lateral-view radiographs; whereas the other 6 cases had a facet joint injury at a visible level on radiographs (Fig. 2).

4. Presence of facet joint injury: agreement

The raw agreement on the presence of a facet joint injury was 72% (κ 0.43; 95% CI, 0.36–0.51) and 77% (κ 0.53; 95% CI, 0.45–0.61) for the first and second plain radiographic survey, respectively. Higher raw agreement values were seen in the third and fourth surveys with CT scan images: 89% (κ 0.75; 95% CI, 0.64–0.86) and 90% (κ 0.76; 95% CI, 0.64–0.88), respectively. Intra-rater reliability was strong for both plain radiography (κ 0.85; 95% CI, 0.72–0.98) and CT scan images (κ 0.92; 95% CI, 0.78–0.1.06) (Table 5).

5. Level of injury

The agreement on determining the most severely affected

level of injury could be calculated for those cases that reached 100% agreement on the presence of a facet joint injury. Where the numbers of cases reaching 100% agreement in the first 2 plain radiographic surveys were low (5 and 7 cases, respectively), inter- and intra-rater agreement coefficients on the level of injury were higher than seen in the CT scan images surveys (Table 5).

6. Injury characteristics

There was poor agreement on whether an injury was of uni- or bilateral nature in the first (κ 0.21; 95% CI, –0.13 to 0.55) and second (κ 0.26; 95% CI, 0.12–0.41) plain radiography surveys. In contrast, the CT scan image surveys resulted in fair to good inter-rater agreement values (Table 5). Intra-rater reliability was strong for both plain radiography (κ 0.87; 95% CI, 0.59–1.14) and CT scan images (κ 0.88; 95% CI, 0.69–1.07).

Similarly, poor agreement values were seen for classifying the injuries according to the injury characteristics listed in Table 2 using plain radiographs (Table 5). CT scan image surveys resulted in higher inter-rater agreement values. Intra-rater reliability was fair to good for plain radiography (κ 0.54; 95% CI, 0.39–0.69) and CT scan images (κ 0.68; 95% CI, 0.59–0.78), respectively.

7. Vertebral body translation

Compared to the plain radiography surveys, inter- and intra-rater agreement on the vertebral body translation was higher for the CT scan image surveys (Table 6).

Discussion

In this radiologic imaging measurement study, more than half of the facet joint injuries detected on CT scan images

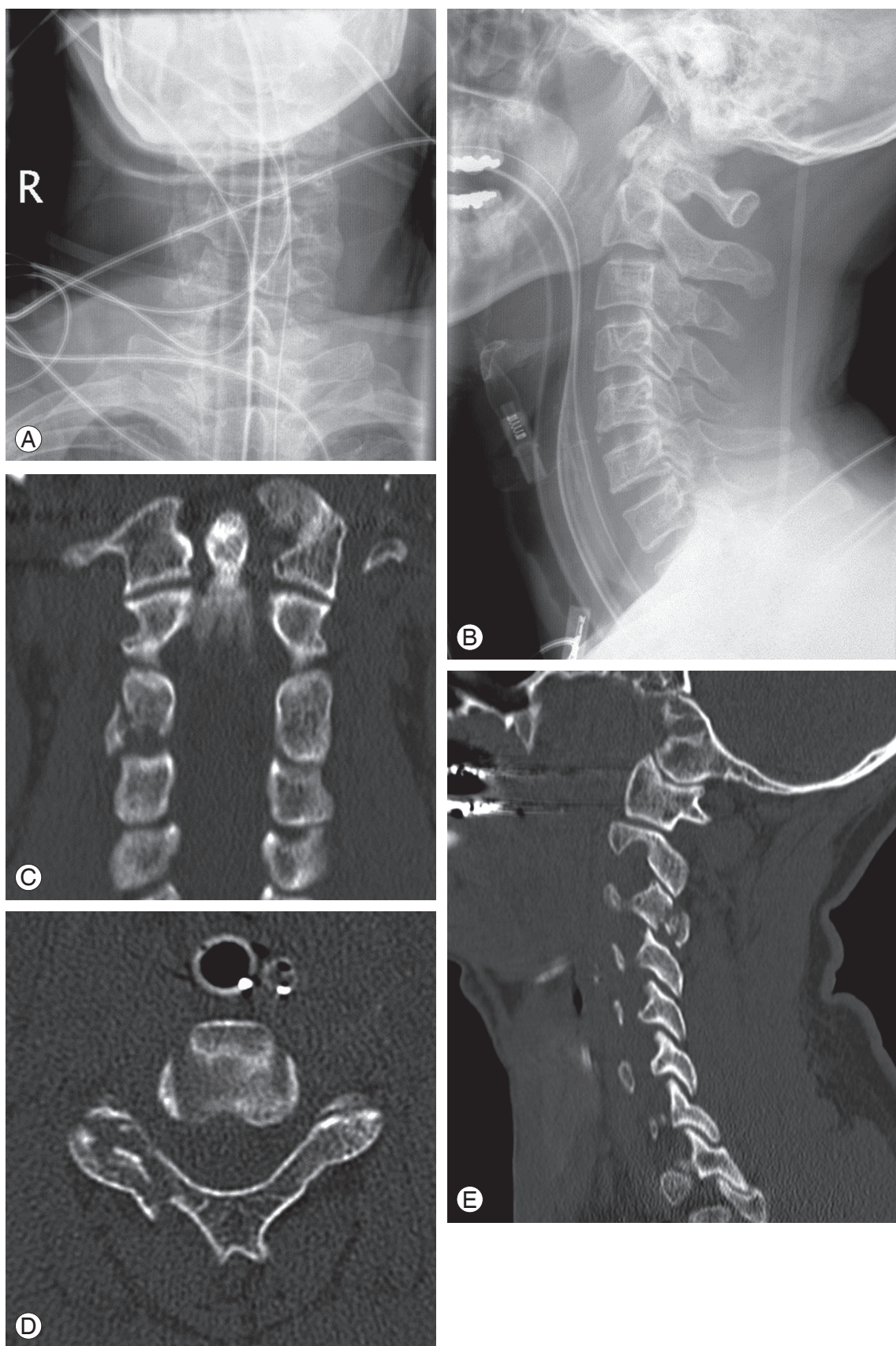


Fig. 2. (A–E) This case represents a 42-year-old man who sustained both a closed head injury and a cervical spine injury.

Table 5. Estimates of inter- and intra-observer reliability

Measure (imaging modality)	Inter-rater reliability			Intra-rater reliability		
	Case	Fleiss Kappa	95% Confidence interval	Case	Cohen's Kappa	95% Confidence interval
Most caudally visible level C-spine						
X-ray						
Test	30	0.54	0.47–0.60	30	0.71	0.61–0.80
Re-test	30	0.55	0.49–0.62	-	-	-
Presence of injury						
X-ray						
Test	30	0.43	0.36–0.51	30	0.85	0.72–0.98
Re-test	30	0.53	0.45–0.61	-	-	-
Computed tomography						
Test	30	0.75	0.64–0.86	30	0.92	0.78–1.06
Re-test	30	0.76	0.64–0.88	-	-	-
Level of injury						
X-ray						
Test	5	1.00	0.86–1.14	5–10	0.95	0.79–1.11
Re-test	7	0.90	0.80–0.99	-	-	-
Computed tomography						
Test	14	0.75	0.66–0.84	10–20	0.86	0.75–0.98
Re-test	12 ^{a)}	0.95	0.87–1.03	-	-	-
Type of injury: uni- or bilateral						
X-ray						
Test	4 ^{b)}	0.21	-0.13 to 0.55	4–10	0.87	0.59–1.14
Re-test	7	0.26	0.12–0.41	-	-	-
Computed tomography						
Test	14	0.58	0.46–0.71	11–20	0.88	0.69–1.07
Re-test	13 ^{b)}	0.54	0.39–0.69	-	-	-
Type of injury: categorisation						
X-ray						
Test	4 ^{b)}	0.23	0.07–0.39	4–10	0.54	0.39–0.69
Re-test	7	0.12	0.02–0.22	-	-	-
Computed tomography						
Test	14	0.47	0.42–0.53	11–20	0.68	0.59–0.78
Re-test	13 ^{b)}	0.34	0.29–0.39	-	-	-

^{a)}Two cases with one missing score; ^{b)}One case with one missing score.

were not detected on plain radiographs of the cervical spine. Since the injured levels were visible on the lateral views of plain radiographs in more than half of the cases, the missed injuries could not be simply attributed to poor visibility of lower cervical spinal levels. When compared to plain radiographs, CT scan images resulted in superior

inter- and intra-rater agreement values for assessing morphologic injury characteristics of facet joint injuries. The fair to good agreement values found for the facet joint injury categories indicate that further refinement of definitions of these injury categories should be considered in future morphological classifications of the subaxial cervi-

Table 6. Estimates of inter- and intra-observer reliability for anterior vertebral translation

Measure (imaging modality)	Inter-rater reliability			Intra-rater reliability		
	Case	ICC	95% CI	Case	ICC	95% CI
Anterior vertebral translation						
X-ray						
Test	5	0.29	0.03–0.81	7–10	0.86	0.75–0.97
Re-test	7	0.30	0.08–0.73	-	-	-
Computed tomography						
Test	14	0.82	0.68–0.92	16–21	0.96	0.94–0.98
Re-test	14	0.72	0.54–0.88	-	-	-

ICC, intraclass correlation coefficient; CI, confidence interval.

cal spine.

The cases under study represent a pseudo-random sample from a consecutive series of high-energy, blunt trauma patients. Since the focus of this study was to examine the accuracy and precision of facet joint injuries of the subaxial cervical spine, we pre-defined a preferred ratio of 2 out of 3 cases with a facet joint injury. As this ratio was not met, 5 cases without facet joint injury were replaced by 5 cases with facet joint injury. This focused approach resulted in a limited epidemiological validity of our findings. Nonetheless, this study concurs with findings from previous studies [8], demonstrating the high specificity (i.e., ratio of true negative cases to all negative cases) of plain radiographs to detect cervical spine injuries. However, from a clinical perspective, the sensitivity of plain radiographs to detect cervical spine injuries is of primary interest. The current study is the first to demonstrate an alarmingly low sensitivity of plain radiographs (42%) to detect facet joint injuries of the cervical spine. Clearly, this study adds to the body of literature advocating the use of CT of the cervical spine in high-energy, blunt trauma patients [7-9].

The outcomes of the current study also have implications for future cervical spinal injury classification studies. Sample numbers were too small to calculate the accuracy of uni- or bilateral facet joint injury detection on plain radiographs. However, we demonstrated that the agreement on whether an injury is of uni- or bilateral nature using plain radiography is much lower than using CT scan images. Therefore the current study reinforces the notion that plain radiographs should not be used for the classification of cervical spine injuries [12]. The categorization of facet joint injuries of the cervical spine presented in Table

2 did not lead to satisfactory agreement values in the CT scan surveys. Further refined categorical descriptions of injury and introductions of graphic representations and example cases may result in higher agreement values. Another factor contributing to ambiguity of facet joint injury categories is that some surgeons consider facet joint injuries with the appearance of a unilateral injury on imaging as a bilateral injury by definition [17]. Some authors have advocated the use of magnetic resonance (MR) imaging for classifying facet joint injuries [18], however, a recent systematic review indicated that MR imaging resulted in poor agreement values for detecting discoligamentous injuries of the thoracolumbar spine [19]. This outcome might also be applicable to cervical spine injuries, hence, requires further study.

Strengths of our study included the prospectively collected data from a well-defined cohort of consecutively recruited trauma patients in a tertiary referral hospital, the use of complete series of plain radiographs and CT scan images and the use of raters blinded to clinical case history. Nonetheless, several potential limitations of our study warrant consideration. Although plain radiographs and CT scan images were blinded, 4 sets of randomly ordered cases were assessed in 1 day. This may have led to both recall bias and fatigue bias. For feasibility reasons, a number of randomly selected cases without a facet joint injury were replaced with cases with a facet joint injury. This has resulted in the limited epidemiological validity of our findings. Four fellowship-trained spine surgeons and 4 spinal surgery fellows scored the cases. The number of cases was too small to perform a subgroup analysis to adjust for raters' experience. Finally, the study findings cannot be generalized to the low-energy trauma population.

Hence, further assessment of the accuracy of X-rays in facet joint injury detection in this particular population, are warranted.

Conclusions

Plain radiographs of the cervical spine are not reliable for the assessment of subaxial cervical spine trauma. Missed injuries cannot be simply attributed to poor visibility of lower cervical spinal levels on the lateral views of plain radiographs as raters also missed facet joint injuries on visible levels. When compared to plain radiographs, CT images resulted in equal or superior inter- and intra-rater agreement values for all studied facet joint related radiologic imaging measurements. Whilst the current study indicates that further refinements are needed, optimization of facet joint injury categories should be considered in future morphological classifications of the subaxial cervical spine.

Conflict of Interest

The AOSpine Asia Pacific Board provided financial support for raters' travel and accommodation expenses to attend the investigators' meeting. No benefits in any form have been or will be received from a commercial party related directly or indirectly to the subject of this manuscript. The authors have no financial interest in the subject of this article. The manuscript submitted does not contain information about medical device(s).

Acknowledgments

The authors thank the Department of Radiology, Radboud University Medical Centre Nijmegen, the Netherlands for the availability of imaging material for the current study. We thank Melissa Anenden for her assistance in data collection and data processing. We acknowledge the AOSpine Asia Pacific Board for providing financial support for raters' travel and accommodation expenses to attend the investigators' meeting. We also thank Mr. Derek Lai for his assistance during the preparations of this study.

References

1. Hasler RM, Exadaktylos AK, Bouamra O, et al. Epidemiology and predictors of cervical spine injury in adult major trauma patients: a multicenter cohort study. *J Trauma Acute Care Surg* 2012;72:975-81.
2. Platzer P, Hauswirth N, Jaindl M, Chatwani S, Vecsei V, Gaebler C. Delayed or missed diagnosis of cervical spine injuries. *J Trauma* 2006;61:150-5.
3. Grossman MD, Reilly PM, Gillett T, Gillett D. National survey of the incidence of cervical spine injury and approach to cervical spine clearance in U.S. trauma centers. *J Trauma* 1999;47:684-90.
4. Thomas J, Rideau AM, Paulson EK, Bisset GS 3rd. Emergency department imaging: current practice. *J Am Coll Radiol* 2008;5:811-6.e2.
5. Sheikh K, Belfi LM, Sharma R, Baad M, Sanelli PC. Evaluation of acute cervical spine imaging based on ACR Appropriateness Criteria(R). *Emerg Radiol* 2012;19:11-7.
6. Lee SH, Sung JK. Unilateral lateral mass-facet fractures with rotational instability: new classification and a review of 39 cases treated conservatively and with single segment anterior fusion. *J Trauma* 2009;66:758-67.
7. Bailitz J, Starr F, Beecroft M, et al. CT should replace three-view radiographs as the initial screening test in patients at high, moderate, and low risk for blunt cervical spine injury: a prospective comparison. *J Trauma* 2009;66:1605-9.
8. Gale SC, Gracias VH, Reilly PM, Schwab CW. The inefficiency of plain radiography to evaluate the cervical spine after blunt trauma. *J Trauma* 2005;59:1121-5.
9. Griffen MM, Frykberg ER, Kerwin AJ, et al. Radiographic clearance of blunt cervical spine injury: plain radiograph or computed tomography scan? *J Trauma* 2003;55:222-6.
10. Dvorak MF, Fisher CG, Aarabi B, et al. Clinical outcomes of 90 isolated unilateral facet fractures, subluxations, and dislocations treated surgically and nonoperatively. *Spine (Phila Pa 1976)* 2007;32:3007-13.
11. Schotanus M, van Middendorp JJ, Hosman AJ. Isolated transverse process fractures of the subaxial cervical spine: a clinically insignificant injury or not?: a prospective, longitudinal analysis in a consecutive high-energy blunt trauma population. *Spine (Phila Pa 1976)* 2010;35:E965-70.
12. van Middendorp JJ, Audige L, Hanson B, Chapman

- JR, Hosman AJ. What should an ideal spinal injury classification system consist of? A methodological review and conceptual proposal for future classifications. *Eur Spine J* 2010;19:1238-49.
13. Bono CM, Vaccaro AR, Fehlings M, et al. Measurement techniques for lower cervical spine injuries: consensus statement of the Spine Trauma Study Group. *Spine (Phila Pa 1976)* 2006;31:603-9.
 14. Fleiss JL. Measuring nominal scale agreement among many raters. *Psychol Bull* 1971;76:378-82.
 15. Fleiss JL. *Statistical methods for rates and proportions*. 2nd ed. New York: Wiley; 1981.
 16. Audige L, Bhandari M, Kellam J. How reliable are reliability studies of fracture classifications? A systematic review of their methodologies. *Acta Orthop Scand* 2004;75:184-94.
 17. Nadeau M, McLachlin SD, Bailey SI, Gurr KR, Dunning CE, Bailey CS. A biomechanical assessment of soft-tissue damage in the cervical spine following a unilateral facet injury. *J Bone Joint Surg Am* 2012; 94:e156.
 18. Vaccaro AR, Madigan L, Schweitzer ME, Flanders AE, Hilibrand AS, Albert TJ. Magnetic resonance imaging analysis of soft tissue disruption after flexion-distraction injuries of the subaxial cervical spine. *Spine (Phila Pa 1976)* 2001;26:1866-72.
 19. van Middendorp JJ, Patel AA, Schuetz M, Joaquim AF. The precision, accuracy and validity of detecting posterior ligamentous complex injuries of the thoracic and lumbar spine: a critical appraisal of the literature. *Eur Spine J* 2013;22:461-74.