ORIGINAL ARTICLE

The Diagnostic Value of an X-ray-based Scoring System for Degeneration of the Cervical Spine: A Reproducibility and Validation Study

Maarten vanEerd iD, MD, PhD, FIPP*,†; Jacob Patijn, MD, PhD†; Daan Loeffen, MD‡; Maarten vanKleef, MD, PhD, FIPP†; Joachim Wildberger, MD, PhD‡

*Department of Anesthesiology and Pain Management, Amphia Ziekenhuis, Breda, The Netherlands; †Department of Anesthesiology and Pain Medicine, Maastricht University Medical Center, Maastricht, The Netherlands; ‡Division of Medical Imaging and Clinical Laboratories, Department of Radiology and Nuclear Medicine, University Medical Centre Maastricht, Maastricht, The Netherlands

■ **Abstract**

Background: In interventional pain medicine, cervical facet joint (CFJ) pain is commonly treated with CFJ denervation techniques, almost automatically assuming degeneration of the CFJs as an important cause of CFJ pain. A standard cervical X-ray is still commonly used in the clinical evaluation of patients suspected for CFJ degeneration. Although degenerative features can be visualized by different radiological imaging techniques, the relation between radiological degenerative features of the cervical spine and pain remains controversial. Paramount in order to estimate the clinical usefulness of a radiological imaging is to establish the reproducibility of the radiological scoring system. A reproducible and clinically feasible diagnostic scoring system was developed to estimate cervical degeneration on standard cervical X-rays.

Materials and Methods: A reproducibility study for the interpretation of degenerative abnormalities on standard cervical X-rays was performed, using a dichotomous outcome (degenerative abnormalities present Yes/No). The estimation of intervertebral disc height loss on standard cervical X-rays was validated with computed tomography (CT) scan measurements.

Results: Five radiological degenerative features on standard cervical X-rays (disc height loss, anterior vertebral osteophytes, posterior vertebral osteophytes, vertebral end plate sclerosis, and uncovertebral osteoarthritis) showed a substantial to excellent reproducibility (kappa value ≥ 0.60). The qualitative definition of disc height loss used in the reproducibility study showed a substantial agreement with the actual measurements of disc height loss on CT scan (kappa value = 0.69).

Conclusion: Subjective judgment of a cervical standard X-ray is a reproducible method to demonstrate degenerative abnormalities of the cervical spine. ■

Key Words: cervical pain, diagnosis, facet joint, X-ray

KEY POINTS/HIGHLIGHTS

- There is no reference standard for the diagnosis degenerative cervical facet joint pain.
- Diagnostic blocks can give an indication for the source of pain but not for the cause of pain.
- Therefore, at the time, the only additive means for the diagnosis degenerative cervical facet joint pain is radiological imaging.
- Standard cervical X-ray is the most commonly used imaging technique in patients suspected of degenerative cervical facet joint pain. The clinical useful interpretation of the images demands definition of the different degenerative features, standardization and a reproducible scoring method.
- Cervical facet joints cannot be judged on standard cervical X-rays because of superposition of the different cervical facet joints. However, other generally accepted radiological, degenerative features such as disc height loss, anterior vertebral osteophytes, posterior vertebral osteophytes, vertebral end plate sclerosis and uncovertebral osteoarthritis, show a substantial inter-rater agreement (kappa values 0.63-0.90) and can therefore be used clinically.
- The qualitative definition of cervical disc height loss on cervical X-ray, being the cervical disc height fitting more than 3 times into the posterior vertebral body height of the vertebra below, shows a substantial agreement (kappa value of 0.69) with the actual measurements on CT scans in the same patients of the defined disc height loss.

INTRODUCTION

In interventional pain medicine, cervical facet joint (CFJ) pain is commonly treated with CFJ denervation techniques, almost automatically assuming degeneration of the CFJs as an important cause of CFJ pain.[1,2] Clinically, the diagnosis of CFJ degeneration is supported by standard cervical X-rays being the most widely available and least expensive imaging technique.[3]

Asymptomatic patients may have degenerative abnormalities and patients with pain may show only minor degenerative changes, indicating the complexity of neck pain.[4–10]

Therefore, the relation between radiological degenerative abnormalities of the cervical spine and pain remains controversial. However, studies evaluating the relation between radiological defined degenerative features and pain used different radiological grading systems to quantify the cervical degenerative changes. Most of these studies used grading systems that were not previously tested for interobserver reproducibility.[5–8,10,11]

Furthermore, in order to define a population of patients with degenerative neck pain, radiological imaging techniques are presently the most important means we have.

Three grading systems for radiographic cervical degenerative abnormalities have been tested for interobserver agreement.[12–15]

These radiographic grading systems show a wide variety in the definition of the different degenerative abnormalities and a varying inter-rater agreement with kappa values ranging from 0.3 to 0.8. Furthermore, most of these radiographic grading systems for cervical degenerative abnormalities are difficult to apply in daily clinical practice because of their complexity. Some of these graded scoring systems sum up the different degenerative abnormalities to an "overall degree of degeneration." However, a summation-score provides no information on which degenerative features are responsible for a high sum score. Therefore, given the differences in the inter-rater agreements between the different degenerative abnormalities, it is recommended to report each of the cervical degenerative features separately.[13]

The most recently proposed radiography-based grading system for cervical degeneration uses the following three degenerative features: (1) height loss of the intervertebral disc, (2) anterior osteophytes, and (3) end plate sclerosis, as well as an overall degree of disc degeneration.[14] These authors report qualitatively and to a certain degree quantitatively the degree of degeneration. The cervical disc height loss was defined as the middle disc height compared to a normal middle disc height at an adjacent level in the same subject.

Degeneration of the CFJs was determined with computed tomography (CT) scan because, according to the authors of this study, the CFJs are better visualized on CT scan.[14]

Anatomic validation of cervical degenerative abnormalities, such as intervertebral disc height, appearance of the disc, and end plate cartilage and osteophyte formation, were evaluated compared to standard X-rays of the specimens only in one study, using 28 human cadavers.[13]

These authors used a macroscopic, anatomic grading system as a reference standard with the highest kappa values (weighted kappa values 0.7–0.9) for the separate variables diffuse sclerosis, osteophyte formation, and disc height loss.

It is assumed in interventional pain medicine that degeneration of the CFJs is an important cause of CFJ pain. Therefore, first reproducibly defining cervical degeneration is paramount.

The aim of this study was to develop a reproducible and clinically feasible diagnostic scoring system for cervical degeneration on standard cervical X-rays.

In addition, the validity of the intervertebral cervical disc height loss on lateral radiographs was estimated by comparing the qualitative estimation of intervertebral disc height loss with the quantitative measurement of the disc height loss on CT scan.

MATERIALS AND METHODS

Study set-up

The study consisted of two parts:

1. Evaluation of the reproducibility of six potential degenerative signs on lateral and anteroposterior cervical X-rays being: cervical disc height loss, anterior vertebral osteophytes, posterior vertebral osteophytes, vertebral end plate sclerosis, facet joint osteoarthritis, and uncovertebral osteoarthritis.[16]
2. Validation of the qualitatively defined cervical disc height loss on lateral X-rays with disc and vertebral height measurements on CT images of the cervical spine in the same subjects. For measurements of cervical disc height and vertebral body height on CT imaging, we used a thin collimated data set with a reasonable overlap. The data were fully reformatted in the 3D postprocessing environment of the scanner. As a result, we were able to reconstruct individual frontal, transversal and sagittal Multi Planner Reformatting (MPR).

Cervical disc height loss was defined as the disc height referred to the posterior vertebral body height.

According to the available literature on cervical spine X-ray imaging, cervical intervertebral disc height loss was defined as the vertebral disc height fitting three times or more in the height of the posterior side of the vertebral body of the level below.[17–19]

Posterior vertebral body height and not anterior vertebral body height was chosen because of the assumption that posterior vertebral body height is more constant during aging compared to anterior vertebral body height. To ascertain this assumption, anterior and posterior vertebral body height was measured in a gender matched population for six age brackets (10–19, 20–29, 30–39, 40–49, 50–59, and 60–69 years).

Reproducibility study

The reproducibility study was performed according to the International Academy for Manual/Musculoskeletal Medicine (IAMM) protocol.[20] This protocol comprises three different standardized phases (training phase, overall agreement phase, and study phase). In the study phase, the P_{index}–50%-method was used to address the problem of the mutual dependency of the prevalence of the index condition with the kappa value. With this method, a low kappa value, due to a too high or too low prevalence of the index condition, is avoided.[21]

Kappa values over 0.60 for a test reflect an interobserver agreement that is considered acceptable to use the test in daily practice and was therefore used as a cutoff point.[22]

A precise definition of measurement of the cervical disc height and the vertebral height on lateral radiographs is described by Frobin et al.[18] In this study, disc height is measured at the anterior side of the disc with a correction for the angle between the adjacent vertebrae. Vertebral height is defined as a dimensionless number by dividing the anterior vertebral height by the mean depth of the vertebra (summation of superior and inferior vertebral depth). With this sophisticated measurement method, not without theoretical flaws (e.g., with respect to the representability of the healthy control group), they found a ratio of angle corrected anterior disc height to vertebral height of 0.35 in a population of healthy subjects (20–45 years, mean = 32 years).

This ratio (0.35) indicates that disc height amounts to roughly one third of the vertebral height. Therefore, we defined disc height loss in the reproducibility study as the middle disc height fitting more than three times in the posterior vertebral body height of the vertebral body below.

Instead of the anterior vertebral height as used by Frobin et al., we used the posterior vertebral body height.[18] The sum of cervical vertebral body wedging of men and women is kyphotic (anterior height of the vertebral body is smaller than the posterior height). The shape of the vertebral bodies will most probably undergo changes with age.[23] We assume a lesser decrease of posterior vertebral body height compared to anterior vertebral body height during aging.[18] To confirm this assumption, we measured in the validation study the anterior and posterior vertebral body height (in mm) in a gender- and age-matched population as well.

Source population. X-ray sets of subjects for the different phases of the reproducibility study were selected out of 8300 cervical X-rays from consecutive 8300 subjects, made at the Emergency Department Maastricht University Medical Centre in 4 consecutive years (Figure 1, source population 1). Enrollment took place, consecutively, on base of date of entry.

For the validation study, X-ray sets of 1180 subjects were selected out of the above-mentioned source of cervical X-rays in which a cervical CT scan was performed at the same day. Included were men and women between the ages of 10–80 years (see Figure 1, source population 2).

Exclusion criteria. Excluded were subjects with fractures of the cervical spine, prior cervical surgery, and congenital abnormalities of the cervical spine.

Medical ethics board approval was obtained (METC 16-4-139). Patients gave informed consent that their medical radiological data can be used for scientific purposes. After selection, the radiological data were anonymized and stored in a database that was only accessible for the two observers of the study.

Reproducibility study

Six different degenerative abnormalities of the cervical spine were evaluated: disc height loss, anterior osteophytes, posterior osteophytes, uncovertebral osteoarthritis ("Suppenteller" Phenomenon),[16] end plate sclerosis, and facet joint osteoarthritis. A separate reproducibility study was planned for each degenerative abnormality.

In the training phases of the reproducibility studies, two observers, both experienced pain specialists (> 20 years), agreed about the definitions and final judgments of a particular cervical degenerative abnormality of the cervical segments C2/C3–C6/C7. A dichotomous judgment for the existence of a particular cervical degenerative abnormality was used (Yes = present / No = not present).

Already in the training phase, it became clear that, due to superposition, the left and right facet joints could not be separately distinguished on lateral X-rays (Figure 2). As a consequence, osteoarthritis of the facet joints was not measured on the X-rays for the subsequent overall agreement and study phase.

The following definitions for the five remaining degenerative abnormalities were used.

Disc height loss (X-ray, lateral view). Performance of test – Disc height is defined as the distance between the two end plates in the middle of the two vertebral bodies of the respective cervical segment (Figure 3a).

Judgment of disc height loss – Disc height loss is considered present when the intervertebral disc height fits more than three times in the height of the posterior side of the vertebral body of the level below. Disc height loss is judged as present or not present.

Anterior osteophytes (X-ray, lateral view). Performance test – Computerized lines are drawn on the lateral X-ray of the depicted cervical spine.

In case of the vertebral body of C2, a line is drawn running along the lower anterior side of the vertebral body of C2 at the ventral side of the boundary of the vertebral body.

In case of the vertebral bodies of C3–C7, a line is drawn along the middle section of the anterior side of the vertebral body and caudally and cranially extended to the superior and inferior end plate of the vertebral body (Figure 3b).

Judgment anterior osteophytes – All bony outgrows of the vertebral body, both at the top and the bottom of the anterior side that are outside the above-mentioned defined line, are considered as anterior osteophytes. Anterior osteophytes are present if one or two bony outgrows are seen.

Posterior osteophytes (X-ray, lateral view). Performance test – Equal to the procedure of defining anterior osteophytes only at the posterior side of the vertebral bodies (Figure 3c).
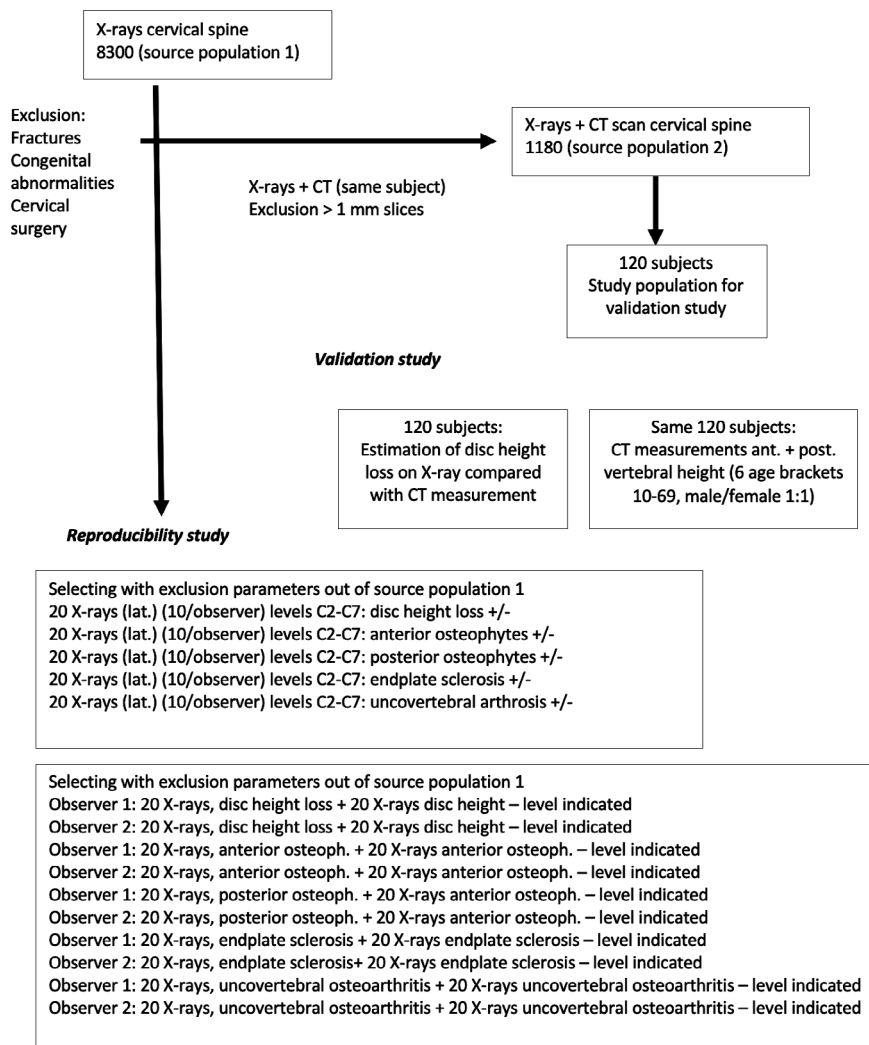
FIGURE 1. Flow chart of selection of radiological images. CT, computed tomography

Judgment anterior osteophytes – All bony outgrowths of the vertebral body, both at the top and the bottom of the posterior side that are outside the above-mentioned defined line, are considered as posterior osteophytes. Posterior osteophytes are present if one or two bony outgrowths are seen.

End plate sclerosis (X-ray, lateral view). *Performance test* – For every cervical segment, the lines of the margins of the constituent vertebral bodies define the end plates (Figure 3d).

Judgment end plate sclerosis – End plate sclerosis is present if one of the above-defined lines has a broader, whiter, and/or a more irregular aspect. Per margin, at least two different aspects must be present.

Uncovertebral osteoarthritis (X-ray, anteroposterior view). *Performance test* – The lateral sides of the top of the vertebral bodies of C3–C7 are judged (Figure 3e).

Judgment uncovertebral osteoarthritis – The margin of the normally cup-shaped configuration of the top of the vertebral body is laterally deflected, resulting in the shape of a soup plate ("Suppenteller" Phenomenon).[16]

The training phase (phase 1) of the test was followed by an overall agreement phase (phase II).

In the overall agreement study, one observer picked 20 X-rays at random out of the source population (source population 1, Figure 1), checked the X-ray data for exclusion criteria, and judged if a degenerative feature was present or not for each cervical segmental level. Then, the second observer judged the same 20 X-rays on

FIGURE 2. X-ray in the lateral view. Note that the cervical facet joint (CFJ) spaces of C2–C3, C3–C4, C4–C5, and C6–C7 are difficult to judge because of superposition of the bilateral facet joints

a separate console and filled in a separate form. This procedure was performed for each degenerative feature (total of 5 × 20 X-rays = 100 X-rays; Figure 1).

In the study phase (phase III), each observer selected 20 positive X-rays and 20 negative X-rays for each of the 5 defined degenerative features out of the source population (source population 1, Figure 1).

After completing the procedure, the forms were collected and matched for statistical analysis.

Blinding procedures. During the test procedures in the overall agreement phase and the study phase, no communication between the two observers was allowed. The radiographs were individually scored each at a separate console, blinded to each other and recorded at two separate forms: one with data number and cervical segmental level and one with data number, cervical segmental level, and judgment (Yes/No). Afterward, the reports were collected for data analysis.

Validation study

Disc height as measured with a CT-multiplanar reformatting imaging technique was used as the reference test

in the validation study. With the 3-D reformatting technique, cut-planes with different colors and therefore access to the frontal, transversal, and sagittal planes at the same time are received. A perpendicular approach was performed by raising the perpendicular line to the depicted segment. The CT disc height measurement technique used was as follows:

1. First, the distance between the left and right medial facet joint margin or the cervical vertebral laminae was measured in the frontal plane and a vertical line was drawn in the middle of this distance (Figure 4a,b).
2. A line was drawn over the superior and inferior vertebral end plates and the perpendicular vertical distance in the middle of these lines was taken as a measure of disc height (mm) in the middle of the disc in the frontal and sagittal planes.
3. Then, in the sagittal plane, both the posterior and anterior vertebral body height of the vertebral body below was measured (mm).
4. The ratio between disc height and posterior vertebral height was calculated.

In order to test the validity of the measurement method, we performed a test-retest of the CT-multiplanar measurement method on 20 CT scans.

For the measurements of anterior and posterior vertebral height of the levels C3–C7 for 6 age brackets (10–19, 20–29, 30–39, 40–49, 50–59, and 60–69 years), 120 subjects (20 subjects per age bracket, male/female ratio 1:1 for each age bracket) were selected out of the source population of 1180 subjects. To assure the quality of the CT images, only CT images with 1 mm slices were included.

For the validation study, the same 120 CT scans out of the source population of 1180 subjects (CT-scan imaging and X-rays of the cervical spine performed on the same day) were selected. Over the years, different multislice CT scanners have been used in our clinic, ranging between 2 and 64 slice CT systems representing 4 different venders and 5 models. We evaluated images of models from Siemens Medical Solutions, Philips Medical Systems, Toshiba Medical Systems, and Picker International from General Electric Co. Ltd. Independently of the model that acquired the images, CT scan slices with 1.0 mm slice thickness with reconstructable overlap of 20%–30% using a sharp reconstruction kernel were included. The scans were performed from the occiput to thoracic vertebra 1 and sagittal and coronal reformations were reconstructed. All images were reviewed with a PACS at 3-megapixel resolution.
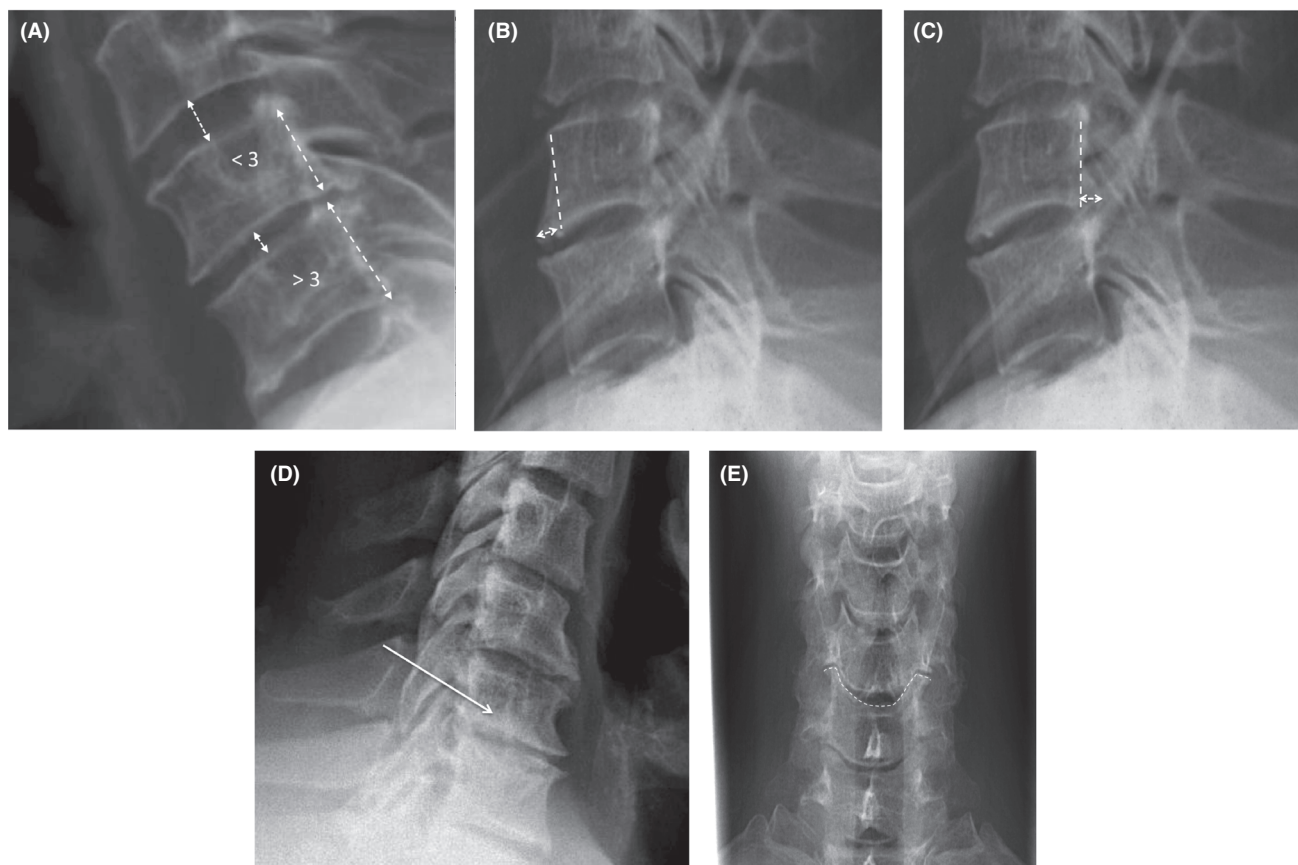
FIGURE 3. (a) Procedure disc height loss (X-ray, lateral view). (b) Procedure anterior osteophytes (X-ray, lateral view). (c) Procedure posterior osteophytes (X-ray, lateral view) double arrow points to the posterior osteophyte. (d) Procedure endplate sclerosis (X-ray, lateral view), the arrow points to end plate sclerosis. (e) Procedure bilateral uncovertebral osteoarthritis, "Suppenteller phenomenon" (X-ray, anteroposterior view)

The scans were reviewed and interpreted by 2 pain specialists with each over 20 years of experience (authors M.v.E. and J.P.). On the X-rays, disc height loss was defined as the mid-intervertebral disc height fitting more than three times in the height of the posterior side of the underlying vertebral body (Yes/No). This is the same definition as used in the reproducibility study.

On CT scan, the ratio of measured posterior vertebral body height and measured disc height was defined positive when the ratio was < 3.5.

On each X-ray and CT scan, all disc levels from C2 to C7 were judged and measured resulting in a potential data of 600 disc height measurements.

Statistical analysis

Kappa values were calculated as a measure for interobserver agreement for the reproducibility of the different degenerative features and for the agreement of the qualitatively based estimation of disc height loss on X-ray and the quantitative determination of disc height loss on CT, together with the prevalence of the index condition (P_{index}), the overall agreement, and confidence intervals (P_{obs}).[21,24]

For the test/retest procedure of the measurements of disc height with MPR, the intraclass correlation coefficient (ICC) for single measurements was calculated.[25]

Proportion statistics were used as a measure of agreement and an overall proportion of agreement with 95% confidence intervals were calculated.[26]

A cutoff value of 0.7 reflecting a high positive (negative) correlation for the Pearson's coefficient was used.[27]

RESULTS

Reproducibility study

In total, 100 radiographs were used for the overall agreement phase (phase II) and 200 radiographs for the
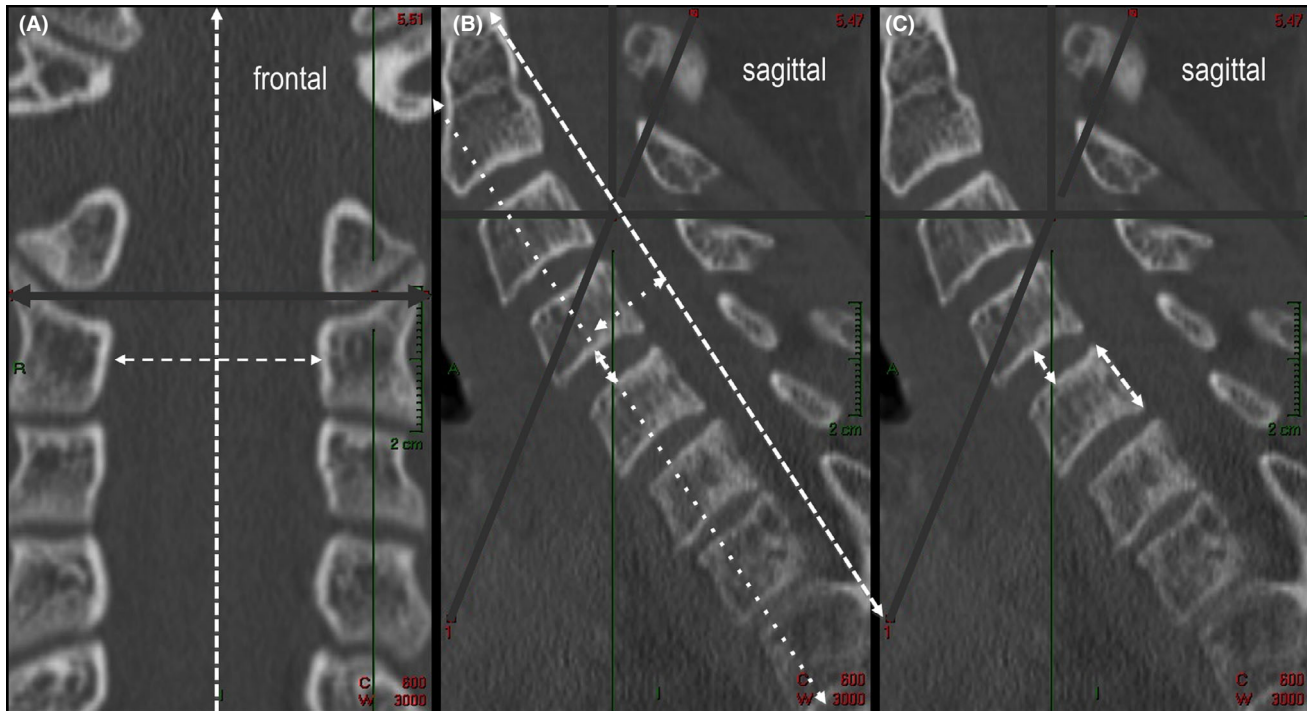
FIGURE 4. (a) In the frontal plane, the vertical line of the sagittal plane is placed in the midline (between both laminae). (b) Subsequently, in the sagittal plane, the vertical line is shifted toward the middle of the vertebral body (dotted line). (c) In the position of the dotted line in Figure 4b, the disc height and posterior vertebral body height (and also anterior vertebral body height was estimated) were measured

study phase (phase III). The agreement between the two observers was almost perfect (above 0.80) for disc height, anterior osteophytes, and end plate sclerosis and substantial (above 0.6) for posterior osteophytes and uncovertebral osteoarthritis.

The kappa values, the overall agreement (P_{obs}) the prevalence of the index condition (P_{index}) for the five degenerative features are listed in Table 1.

TABLE 1. Kappa value, overall agreement (P_{obs}) and prevalence of the index condition (P_{index}) for the five degenerative features

	Kappa value with CIs	Overall agreement (P_{obs})	Prevalence index condition (P_{index})
Disc height loss	0.85 (CI = 0.69–1.01)	0.93	0.54
Anterior osteophytes	0.85 (CI = 0.68–1.01)	0.93	0.51
Posterior osteophytes	0.63 (CI = 0.38–0.88)	0.83	0.61
Uncovertebral osteoarthritis	0.75 (CI = 0.54–0.95)	0.88	0.46
End plate sclerosis	0.90 (CI = 0.76–1.04)	0.95	0.48

Abbreviation: CI, confidence interval.

Validation study

The test-retest of the CT multislice measurement method for intervertebral disc height showed an ICC of 0.93, showing an excellent reproducibility of the described CT measurement method for intervertebral disc height.

For the posterior and anterior vertebral body heights, the ICCs were, respectively, 0.98 and 0.98. Both posterior and anterior vertebral body heights showed the same significant correlation ($p < 0.000$) with the age group. Compared to the posterior vertebral body height, there was a tendency of a more pronounced decrease in the anterior vertebral body height compared to the posterior vertebral body height during aging.

For the validity of the disc height, the agreement between the subjective ratio of the disc height to the posterior vertebral body height on lateral radiographs of the cervical spine and disc height to posterior vertebral body height as measured with multislice CT scan was calculated and expressed in the kappa value. On the 120 X-rays, 104 disc-levels were not possible to judge, because the lower cervical segmental levels are sometimes not clearly visible on lateral X-rays due to

TABLE 2. The 2 × 2 contingency table for validation of disc height loss

	Quantitative (CT) measurement ratio disc/vertebral height <0.35		
	Yes	**No**	
Qualitative judgement (X-ray) **Yes**	30	12	42
No	12	442	454
	42	454	496

Qualitatively: 3× disc/posterior vertebral height and quantitatively disc <3.5 posterior vertebral height.

shoulder over-projection, resulting in a total of 496 judgments and measurements (Table 2). For the CT-measurements, a ratio of > 0.35 was used to define disc height loss. This ratio choice is based on the results of a previous study in which the ratio ranged from 0.30 to 0.39.[18]

A kappa value of 0.69 was found with a P_{obs} of 0.95 and a P_{index} of 0.08 (Table 2).

DISCUSSION

The inter-rater agreement (reproducibility) of a dichotomous scoring system for cervical degenerative abnormalities on plain radiography was measured. The radiological degenerative abnormalities, disc height loss, anterior vertebral osteophytes, posterior vertebral osteophytes, vertebral end plate sclerosis, and uncovertebral osteoarthritis showed a substantial to excellent reproducibility (kappa values = 0.63–0.90).

An important finding with respect to interventional pain strategies aimed at the CFJs is that the CFJs cannot properly be visualized on lateral radiographs due to superposition of the different depicted CFJs (see Figure 1).

Only three existing graded scoring systems for the assessment of cervical degeneration have previously been tested for interobserver agreement.[14,15,28,29] Although the five degenerative features defined in our study have similarities with the last published graded scoring system,[14] we added the item posterior vertebral osteophytes[29] and the item uncovertebral osteoarthritis. Uncovertebral osteoarthritis was included because the uncovertebral articulations are common sites for osteoarthritic changes and a potential pain generator in the cervical spine.[30,31]

TABLE 3. Measurement of anterior and posterior vertebral body height on CT-MPR for each vertebral level (120 subjects, 20 per age bracket male/female 1:1)

	Anterior vertebral body height mm decrease/10 year (mm)	Posterior vertebral body height mm decrease/10 year (mm)
C_3	0.3	0.20
C_4	0.2	0.04
C_5	0.3	0.07
C_6	0.2	0.02
C_7	0.1	0.04

Abbreviation: CT-MPR, computed tomography multi planner reformatting.

Second, we did not define disc height loss,[14] as disc height loss compared with other cervical levels in the same subject but as a ratio of disc height to height of the posterior side of the vertebral body below. The ratio between disc height and vertebral body height (mean depth or anterior vertebral height) in a normal population (mean age = 32 years, SD = 11) is roughly one third (0.35).[18] We defined disc height loss qualitatively as the cervical disc height fitting more than three times into the posterior vertebral height of the vertebra below.[17,18] The posterior vertebral body height was chosen because the posterior vertebral body height changes less with age (see Table 3).

Third, we did not use a graded scoring system. Graded scoring systems can be useful in longitudinal studies to study progression of cervical degeneration. However, a dichotomous scoring system is more suitable to answer the question if cervical spine degeneration is present or not.

As mentioned above, one study defined disc height loss qualitatively by comparing the middle disc height with the middle disc height of the adjacent level in the same subject.[14] They report a somewhat lower ICC of 0.728 (95% confidence interval [CI] = 0.54–0.86)

compared to our study (ICC = 0.85, 95% CI = 0.69–1.0). However, a major problem of this definition of disc height loss is that it is dependent on the assumption that the disc height of the adjacent cervical level is a normal disc height.

Another reproducibility study defined disc height loss as the sum of the anterior and posterior disc height compared to the respective individual disc heights before degeneration based on a set of normal values.[13] The interobserver agreement of disc height loss with a kappa value of 0.83 reported in this study is comparable with our reported kappa value for interobserver agreement of disc height loss (kappa value = 0.85).[13] However, our qualitative definition of cervical disc height loss is less complex and, in our opinion, can more easily be used in daily clinical practice in interventional pain management.

End plate sclerosis of the cervical vertebra or subchondral bone thickening is radiologically defined on a standard cervical X-ray by broadening and whitening of the vertebral end plate line.[13] We report an excellent interobserver agreement for end plate sclerosis with a kappa value of 0.9 compared with kappa values of 0.71 and 0.31 in the previous studies.[13,14] A possible explanation for the higher kappa value found in our study is that we used a dichotomous judgment instead of a more complex graded scoring method.

It is reported that degenerative abnormalities of the CFJs are difficult to grade because of often poor visualization of the CFJs on a standard lateral X-ray.[15] Therefore, cervical CT imaging was recommended.[14]

In our study, it became apparent that proper judgement of CFJ degeneration on a lateral standard cervical X-ray is not possible.

Although a cervical CT scan may be used in the future to demonstrate degenerative abnormalities, the current pain management practice is that standard cervical X-rays are routinely made in patients with neck pain.

A limitation of our study, using a standard cervical X-ray, is the fact that the CFJs cannot be judged for degeneration because of superposition of the CFJs in lateral X-rays. However, if the other cervical degenerative abnormalities evaluated in our study are caused by mechanical loading, this may also affect the CFJs and, subsequently, lead to signs of facet joint osteoarthritis.[32–43] As such, the judgment of standard cervical X-rays can still be used to demonstrate cervical degenerative abnormalities in a reproducible way.

There is an absence of definitions and reproducible judgment procedures for degenerative abnormalities on cervical CT scans. Therefore, no validation could be performed for the degenerative abnormalities of the anterior vertebral osteophytes, posterior vertebral osteophytes, vertebral end plate sclerosis, and uncovertebral osteoarthritis. Only validation of our qualitative definition of disc height loss was possible with standardized measurements of cervical disc height on a CT scan.

Another limitation of our study is that most of the cervical X-rays were performed in a weight-bearing condition and all cervical CT images were performed in a nonweight-bearing condition. The literature about differences in cervical disc morphology in standing versus the supine position is sparse.[44] However, a higher disc height in the supine position on a CT scan compared to the disc height on cervical X-rays might have influenced the agreement measurement in the validation study.

To answer the clinical question whether degenerative abnormalities of cervical spine degeneration are present or not, we studied the reproducibility of a dichotomous radiographic scoring system for cervical degenerative abnormalities on standard cervical X-rays. In the absence of a reference standard for degenerative CFJ pain, we defined in a study on interventions for degenerative CFJ pain that at least three of the five described degenerative features had to be present assuming an interdependent association between CFJ degeneration and other radiographic cervical degenerative abnormalities.[45] This assumption has to be substantiated with future research.

CONFLICT OF INTEREST

None of the authors have a conflict of interest.

REFERENCES

1. Bykowski JL, Wong WH. Role of facet joints in spine pain and image-guided treatment: a review. *AJNR Am J Neuroradiol*. 2012;33:1419–26.

2. Manchikanti L, Kaye AD, Boswell MV, et al. A systematic review and best evidence synthesis of the effectiveness of therapeutic facet joint interventions in managing chronic spinal pain. *Pain Physician*. 2015;18:E535–82.

3. Bijlsma JW, Berenbaum F, Lafeber FP. Osteoarthritis: an update with relevance for clinical practice. *Lancet*. 2011;377:2115–26.

4. Bogduk N. Degenerative joint disease of the spine. *Radiol Clin North Am*. 2012;50:613–28.

5. Friedenberg ZB, Miller WT. Degenerative disc disease of the cervical spine. *J Bone Joint Surg*. 1963;45:1171–8.

6. Gore DR. Roentgenographic findings in the cervical spine in asymptomatic persons: a ten-year follow-up. *Spine (Phila Pa 1976)*. 2001;26:2463–6.

7. Marchiori DM, Henderson CN. A cross-sectional study correlating cervical radiographic degenerative findings to pain and disability. *Spine (Phila Pa 1976)*. 1996;21:2747–51.

8. Gore DR, Sepic SB, Gardner GM, Murray MP. Neck pain: a long-term follow-up of 205 patients. *Spine (Phila Pa 1976)*. 1987;12:1–5.

9. Peterson C, Bolton J, Wood AR, Humphreys BK. A cross-sectional study correlating degeneration of the cervical spine with disability and pain in United Kingdom patients. *Spine (Phila Pa 1976)*. 2003;28:129–33.

10. Heller CA, Stanley P, Lewis-Jones B, Heller RF. Value of X ray examinations of the cervical spine. *Br Med J (Clin Res Ed)*. 1983;287:1276–8.

11. Stone LS. Joint degeneration and chronic pain: still looking for the missing link. *Pain*. 2009;141:185–6.

12. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. *Ann Rheum Dis*. 1957;16:494–502.

13. Kettler A, Rohlmann F, Neidlinger-Wilke C, Werner K, Claes L, Wilke HJ. Validity and interobserver agreement of a new radiographic grading system for intervertebral disc degeneration: Part II. Cervical spine. *Eur Spine J*. 2006;15:732–41.

14. Walraevens J, Liu B, Vander Sloten J, Goffin J, et al. Qualitative and quantitative assessment of degeneration of cervical intervertebral discs and facet joints. *Eur Spine J*. 2009;18:358–69.

15. Cote P, Cassidy JD, Yong-Hing K, Sibley J, Loewy J. Apophysial joint degeneration, disc degeneration, and sagittal curve of the cervical spine. Can they be measured reliably on radiographs? *Spine (Phila Pa 1976)*. 1997;22:859–64.

16. Aufdermauer M. *Die Spondylosis Cervicalis*. Stuttgart: Hippokrates-Verlag; 1960.

17. Chevrot A, Drape JL, Godefroy D, Dupont AM. [Imaging of the painful cervical spine]. *J Radiol*. 2003;84:181–239.

18. Frobin W, Leivseth G, Biggemann M, Brinckmann P. Vertebral height, disc height, posteroanterior displacement and dens-atlas gap in the cervical spine: precision measurement protocol and normal data. *Clin Biomech*. 2002;17:423–31.

19. Gilad I, Nissan M. A study of vertebra and disc geometric relations of the human cervical and lumbar spine. *Spine (Phila Pa 1976)*. 1986;11:154–7.

20. Patijn J. Reproducibility protocol for diagnostic procedures in Manual/Musculoskeletal Medicine. *Manuelle Medizin*. 2019;57:451–479.

21. Gjorup T. The kappa coefficient and the prevalence of a diagnosis. *Methods Inf Med*. 1988;27:184–6.

22. Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics*. 1977;33:363–74.

23. Been E, Shefi S, Soudack M. Cervical lordosis: the effect of age and gender. *Spine J*. 2017;17:880–8.

24. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Measur*. 1960;20:37–46.

25. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1:307–10.

26. Peacock J, Kerry S. *Presenting medical statistics from proposal to publication*. Oxford: Oxford University Press; 2007.

27. Mukaka MM. Statistics corner: a guide to appropriate use of correlation coefficient in medical research. *Malawi Med J*. 2012;24:69–71.

28. Kellgren JH, Lawrence JS. Radiological assessment of rheumatoid arthritis. *Ann Rheum Dis*. 1957;16:485–93.

29. Kettler A, Wilke HJ. Review of existing grading systems for cervical or lumbar disc and facet joint degeneration. *Eur Spine J*. 2006;15:705–18.

30. Brismee JM, Sizer PS Jr, Dedrick GS, Sawyer BG, Smith MP. Immunohistochemical and histological study of human uncovertebral joints: a preliminary investigation. *Spine (Phila Pa 1976)*. 2009;34:1257–63.

31. Hartman J. Anatomy and clinical significance of the uncinate process and uncovertebral joint: a comprehensive review. *Clin Anat*. 2014;27:431–40.

32. Butler D, Trafimow JH, Andersson GB, McNeill TW, Huckman MS. Discs degenerate before facets. *Spine (Phila Pa 1976)*. 1990;15:111–3.

33. Kirkaldy-Willis WH, Wedge JH, Yong-Hing K, Reilly J. Pathology and pathogenesis of lumbar spondylosis and stenosis. *Spine (Phila Pa 1976)*. 1978;3:319–28.

34. Gries NC, Berlemann U, Moore RJ, Vernon-Roberts B. Early histologic changes in lower lumbar discs and facet joints and their correlation. *Eur Spine J*. 2000;9:23–9.

35. Eubanks JD, Lee MJ, Cassinelli E, Ahn NU. Does lumbar facet arthrosis precede disc degeneration? A post-mortem study. *Clin Orthop Relat Res*. 2007;464:184–9.

36. Fujiwara A, Tamai K, Yamato M, An HS, Yoshida H, Saotome K, et al. The relationship between facet joint osteoarthritis and disc degeneration of the lumbar spine: an MRI study. *Eur Spine J*. 1999;8:396–401.

37. Haig AJ. Paraspinal denervation and the spinal degenerative cascade. *Spine J*. 2002;2:372–80.

38. Kettler A, Werner K, Wilke HJ. Morphological changes of cervical facet joints in elderly individuals. *Eur Spine J*. 2007;16:987–92.

39. Lee MJ, Dettori JR, Standaert CJ, Brodt ED, Chapman JR. The natural history of degeneration of the lumbar and cervical spines: a systematic review. *Spine (Phila Pa 1976)*. 2012;37:S18–30.

40. Suri P, Miyakoshi A, Hunter DJ, Jarvik JG, Rainville J, Guermazi A, et al. Does lumbar spinal degeneration begin with the anterior structures? A study of the observed epidemiology

in a community-based population. *BMC Musculoskelet Disord*. 2011;12:202.

41. Swanepoel MW, Adams LM, Smeathers JE. Human lumbar apophyseal joint damage and intervertebral disc degeneration. *Ann Rheum Dis*. 1995;54:182–8.

42. Vernon-Roberts B, Pirie CJ. Degenerative changes in the intervertebral discs of the lumbar spine and their sequelae. *Rheumatol Rehabil*. 1977;16:13–21.

43. Wilder FV, Fahlman L, Donnelly R. Radiographic cervical spine osteoarthritis progression rates: a longitudinal assessment. *Rheumatol Int*. 2011;31:45–8.

44. Yoganandan N, Umale S, Stemper B, Snyder B. Fatigue responses of the human cervical spine intervertebral discs. *J Mech Behav Biomed Mater*. 2017;69:30–8.

45. van Eerd M, de Meij N, Kessels A, Patijn J, Weber W, Wintraecken V, et al. Efficacy and long-term effect of radiofrequency denervation in patients with clinically diagnosed cervical facet joint pain: a double-blind randomized controlled trial. *Spine (Phila Pa 1976)*. 2021;46:285–93.