Illusions of fusions: Assessing cervical vertebral fusion on lateral cephalograms, multidetector computed tomographs, and cone-beam computed tomographs

Raphael Patcas, a Dominika Tausch, b Nikolaos Pandis, c Mirjana Manestar, d Oliver Ullrich, e Christoph A. Karlo, f Timo Peltomäki, g and Christian J. Kellenberger h
Zurich and Bern, Switzerland, Corfu, Greece, and Tampere, Finland

Introduction: The aims of this study were to compare lateral cephalograms with other radiologic methods for diagnosing suspected fusions of the cervical spine and to validate the assessment of congenital fusions and osteoarthritic changes against the anatomic truth. Methods: Four cadaver heads were selected with fusion of vertebrae C2 and C3 seen on a lateral cephalogram. Multidetector computed tomography (MDCT) and cone-beam computed tomography (CBCT) were performed and assessed by 5 general radiologists and 5 oral radiologists, respectively. Vertebrae C2 and C3 were examined for osseous fusions, and the left and right facet joints were diagnosed for osteoarthritis. Subsequently, the C2 and C3 were macerated and appraised by a pathologist. Descriptive analysis was performed, and interrater agreements between and within the groups were computed. Results: All macerated specimens showed osteoarthritic findings of varying degrees, but no congenital bony fusion. All observers agreed that no fusion was found on MDCT or CBCT. They disagreed on the prevalence of osteoarthritic deformities (general radiologists/MDCT, 100%; oral radiologists/CBCT, 93.3%) and joint space assessment in the facet joints (kappa = 0.452). The agreement within the rater groups differed considerably (general radiologists/MDCT, kappa = 0.612; oral radiologists/CBCT, kappa = 0.240). Conclusions: Lateral cephalograms do not provide dependable data to assess the cervical spine for fusions and cause false-positive detections. Both MDCT interpreted by general radiologists and CBCT interpreted by oral radiologists are reliable methods to exclude potential fusions. Degenerative osteoarthritic changes are diagnosed more accurately and consistently by general radiologists evaluating MDCT. (Am J Orthod Dentofacial Orthop 2013;143:213-20)

In recent years, orthodontists have expressed increasing interest in assessing the cervical spine on a lateral cephalogram. One clinical purpose is the determination of skeletal age based on the association between age-related morphologic changes of the upper cervical vertebrae and the somatic growth curve. 1, 2 A further intent is the evaluation of the craniovertebral angulation to characterize head posture, which has been linked to nasorespiratory function 3 and craniofacial morphology. 4 Moreover, the use of lateral cephalograms has also been recommended to study congenital anomalies of the cervical vertebrae, because cervical vertebral anomalies, particularly fusions, could be related to certain craniofacial syndromes and other dentoskeletal malformations. 5-16 Awareness that the spine is of clinical interest has led to the recommendation to use cephalometric radiographs to routinely screen the cervical vertebrae for anomalies and even to develop a tracing technique of this region. 17
Fusions are most common between the facet joints of the second and third vertebrae (C2 and C3; Fig 1). Like all other cervical vertebral anomalies, osseous fusions are usually asymptomatic and considered to be coincidental findings with no clinical relevance. However, in a few patients, cervical vertebral anomalies cause compression of neurologic structures or biomechanical instability, leading to chronic pain. Associations between cervical vertebral anomalies, notably fusion of C2 and C3, and congenital disorders or dentoskeletal malocclusions have been studied extensively. They include syndromic and nonsyndromic anomalies such as fetal alcohol syndrome and cleft lip and palate. In recent research examining cervical vertebral anomalies on lateral cephalograms, a high prevalence of cervical vertebral anomalies, particularly fusions of C2 and C3, was reported in orthodontic surgical patients with severe skeletal malocclusions. The described associations between cephalometric measurements and fusions include skeletal Class III and mandibular overjet with 61.4% fusions, skeletal deepbite with 41.5% fusions, skeletal open bite with 42.1% fusions, and skeletal Class II and maxillary overjet with 28% and 52.9% fusions, respectively. A similarly high prevalence of fusions has been documented in subjects with condylar hypoplasia with 72.7% fusions (45% in C2 and C3) and in patients with obstructive sleep apnea (46%).

These findings have been challenged by some who argued that it was difficult to reliably determine cervical vertebral anomalies on 1 lateral cephalogram. Considerably lower prevalence numbers (<0.9%) have been reported in other studies with normal populations; this could be because patients with severe malocclusions are significantly different from a normal population, but this dissimilarity in prevalence certainly raises the question as to whether lateral cephalograms are a reliable tool to assess cervical vertebral anomalies. Koletsis and Halazonetis stated that no study investigating the reliability of cephalometric radiography in the cervical region has been published to date. To validate the assessment of the spine on lateral cephalograms, 3-dimensional radiological data and direct observation (on autopsy material) have been suggested. A cadaver study would allow for direct comparisons of different assessment methods and validate each diagnostic approach against the anatomic truth.

In addition, a cadaver study would serve another purpose: diagnostic thinking efficacy evaluates whether the information retrieved from radiologic images leads to a change in the clinician's diagnostic thinking. This efficacy has been evaluated for cone-beam computed tomography (CBCT) in relation to impacted third molars, impacted canines, root resorption of adjacent teeth, and the temporomandibular joint, but it has not been appraised for the cervical spine. Since the cervical spine is a region of interest for the orthodontist, it would be beneficial to assess the diagnostic efficacy of oral radiologists examining CBCT data and to compare it with that of general radiologists analyzing multidetector computed tomography (MDCT) data, and to verify the results against the anatomic findings.

The objectives of this cadaver study were therefore (1) to ascertain whether fusions of C2 and C3 suspected on lateral cephalograms would also be diagnosed by general radiologists on MDCT or oral radiologists on CBCT, and (2) to validate MDCT and CBCT assessments of congenital fusions and osteoarthritic changes against the anatomic truth.

MATERIAL AND METHODS

From a larger collection of perfused cadaver heads, 8 specimens were selected for which analog, postmortem lateral cephalograms were available (tube voltage, 67; tube current, 250 mA; exposure time, 0.04 second; tube current time product, 10 mAs; focus to coronal plane distance, 200 cm). The cadaver heads were supplied by the Institute of Anatomy at the University of Zurich in Switzerland in accordance with state and federal regulations (ie, voluntary body donation program on the basis of informed consent), the Convention on Human Rights and Medicine, and the recommendation of the National Academy of Medical Science. The perfusion was carried out within 4 days after death with a fixation liquid consisting of 2 parts alcohol (70%),
1 part glycerine, and 2% almudor (containing 8.1% formaldehyde, 10% glyoxal, and 3.7% glutaraldehyde). The lateral cephalograms were screened and assessed for potential fusions of cervical vertebrae by an author (D.B.) following the method prescribed in the literature: fusions were identified as an osseous continuity between C2 and C3 without complete separation at the articular facets or intervertebral disc space (see Fig 2 for an excluded specimen with a continuous radiolucent area).7,20,21,35,36 Four specimens (3 female, 1 male; age range, 65-87 years; mean age, 78 years) fulfilled the inclusion criterion of a suspected cervical spine fusion at the level of C2 and C3 and were used for the study.

MDCT was performed on a 40-detector row computed tomography system (Brilliance CT 40; Philips Healthcare, Eindhoven, The Netherlands) with the following scan parameters kept identical for all specimens: tube voltage, 120 kV; tube current time product, 70 mAs; slice collimation, 0.625 mm; pitch, 0.68; reconstruction slice thickness, 0.67 mm; reconstruction increment, 0.33 mm; window level setting, 2000/500 Hounsfield units; voxel sizes, 0.39 mm (x-axis), 0.39 mm (y-axis), and 0.67 mm (z-axis). Sagittal and coronal reformatted images (slice thickness, 1 mm; increment, 0.5 mm) were viewed on a high-resolution diagnostic workstation (dx IDS5; Sectra PACS, Linköping, Sweden).

The CBCT scans were made on a scanner with an amorphous silicon flat panel (KaVo 3D exam; KaVo Dental, Bismarckring, Germany). The following scan parameters were kept identical during all CBCT examinations: tube voltage, 120 kV; tube current time product, 37.07 mAs; reconstruction thickness, 0.25 mm; reconstruction increment, 0.25 mm; voxel size, 0.25 mm (x-axis), 0.25 mm (y-axis), and 0.25 mm (z-axis). Digital imaging and communications in medicine (DICOM) files were reformatted in multplanar reconstructions by using open-source postprocessing software (Workstation version 2.0.5P1; ClearCanvas, Toronto, Ontario, Canada).

Five general radiologists were asked to evaluate the MDCT data, and 5 dentists with special postgraduate training in oral radiology were asked to evaluate the CBCT data. Three areas were assessed for a potential congenital fusion: facet joint (C2-C3) of the left and right articular processes and the intervertebral disc space between the 2 bodies, C2 and C3. Additionally, the raters were requested to perform their radiological appraisals for the left and right facet joints in the following manner: (1) normal joint or (2) osteoarthritis (joint space entirely preserved, partially preserved, or not visible).

All radiologists assessed the images independently, in blinded fashion, and without knowledge of the anatomic findings.

After image acquisition, the cervical spines were isolated en bloc from the cadaver heads. Lipids were dissolved in a Supralan UF solution (Bauer Handels, Fehraltorf, Switzerland) with sodium chloride. Enzymatic maceration was performed with papain (Bauer Handels) at pH 6 to 7 in Supralan UF and a solution containing sodium chloride for up to 14 days. The macerated spines were subsequently analyzed for fusions and osteoarthritis by a board-certified pathologist.

Statistical analysis

A standard statistical software package (version 11.4.1.0; MedCalc Software, Mariakerke, Belgium) was used for the descriptive data analysis. An unweighted Cohen kappa test was computed to evaluate the agreement between the CBCT and MDCT methods.37 To determine interobserver agreement between the 5 CBCT radiologists and the 5 MDCT radiologists, a Fleiss kappa test38 for multiple raters was calculated with StatTools.39

RESULTS

After the enzymatic maceration, the vertebral bodies C2 and C3 could be completely mobilized, proving the absence of congenital bony fusions in these vertebral segments (Figs 3-5). All facet joints showed degenerative osteoarthritic changes including osteophytes, peripheral eburnation, and gross irregularities of the subchondral joint surfaces of varying degrees. The facets of 2 specimens were more severely affected, exhibiting extensive osteophytes and ragged bony joint surfaces (Fig 5).
All raters agreed that no congenital fusion was found by MDCT or CBCT, but there was disagreement concerning the prevalence of the osteoarthritic deformities. General radiologists assessing the MDCT recognized osteoarthritic changes in 100% (40/40) of the joint assessments, and oral radiologists evaluating the CBCT found osteoarthritic changes in 93.3% (38/40). Moreover, when evaluating the narrowing of the joint space in the affected osteoarthritic joints, the 2 rater groups differed substantially (Table I). The concordance...
between the 2 rater groups was rather modest (70%) with a kappa value of 0.452 (SE, 0.132; 95% confidence interval [CI], 0.193-0.711), indicating moderate agreement.40

In addition, there was a considerable difference regarding the agreement within each rater group when evaluating the narrowing of the joint space. The general radiologists assessing the MDCT data agreed more consistently with each other (kappa = 0.612) than did the oral radiologists assessing the CBCT data (kappa = 0.240; Table II). According to Landis and Koch,40 the kappa value of 0.240 for the CBCT/oral radiologists corresponds to fair agreement, and the kappa value of 0.612 for the MDCT/general radiologists denotes substantial agreement.

**DISCUSSION**

The evaluation of congenital vertebral fusions on lateral cephalograms has been studied extensively in the orthodontic literature, associating fusions with diverse anomalies and malocclusions. The use of lateral cephalograms is valuable in the assessment of developmental anomalies of the craniofacial skeleton. The present study demonstrated that the agreement between the two rater groups was moderate, with a kappa value of 0.452, indicating that the radiologists’ assessments were somewhat consistent but not highly reliable. This finding highlights the potential for variability in interpreting these images, which is important for clinicians to consider when evaluating cases of craniofacial dysmorphology.
cephalograms, however, has been challenged by those who argue that 2-dimensional radiographs can yield deceptive impressions of “pseudo-fusions” in the C2-C3 facet joint because of their oblique orientation. This study corroborates this concern by demonstrating that cephalograms do not provide reliable data to assess vertebral fusions in the cervical spine. All 4 specimens assessed positively for fusions on lateral cephalograms proved to be false positives. None had a fusion. Hence, the absence of a continuous radiolucent area between the articular processes (on cephalograms) as the sole radiologic criterion might not be a valid method to identify fusions on 1 lateral cephalogram.

All evaluated joints had osteoarthritic changes, some with gross irregularities and narrowed joint spaces. Based on our findings, a further reason for the erroneous assessment of fusions might be the misinterpretation of osteoarthritic changes as fusions. It is evident that a continuous radiolucent area might fade away because of irregularities, as shown in Figures 5 and 6 (see specimen 3).

The results demonstrate the limitations of lateral cephalograms as a diagnostic tool to assess the spine and raise doubts about the necessity of exposing the cervical vertebrae to radiation and, with it, the thyroid. Hence, the clinical recommendation to apply a neck

**Fig 6.** Coronal *(left)* and sagittal *(right)* reformatted MDCT images *(top)* and CBCT images *(bottom)* of specimen 3. The *purple arrows* point to the irregularly narrowed facet joints C2-C3 with subchondral sclerosis and spondylophytes, but no bony fusion. The images were rotated and cropped to facilitate a direct comparison. Note the close proximity to the edge of the volume in CBCT, seen on the coronal slide.
shield consistently on lateral cephalograms ought to be reinforced. Recent studies demonstrate that, if skeletal age estimation is needed, radiation exposure can be minimized by applying a neck shield and performing an additional hand-wrist radiograph.41

In this study, we also compared the different assessment outcomes between the general radiologists evaluating MDCT and the oral radiologists appraising CBCT of the cervical spine (Fig 6). The results demonstrate 2 important findings: (1) both rater groups performed equally well regarding the exclusion of possible fusions; and (2) concerning the appraisal of osteoarthritic deformities, the general radiologists assessing MDCT performed uncontestably better. They diagnosed osteoarthritic changes correctly in 100% of the cases and did so with considerable consistency in regard to their assessment of the joint spaces. Conversely, oral radiologists evaluating CBCT diagnosed only 93.3% of the osteoarthritic cases correctly and did so with more disagreement among themselves in their assessments of the joint spaces.

Two possible assumptions might explain why oral radiologists evaluating CBCT data do not perform as well as general radiologists with MDCT data. On one hand, oral radiologists are not used to assessing joints. The only joint in the maxillofacial region is the temporomandibular joint, which differs remarkably from other joints. Hence, it could be argued that general radiologists probably perform better because of their broader experience in assessment of articulo-osseous pathologies. On the other hand, there is an inherent problem with CBCT data. The image quality in the midplane is superior to more peripheral regions because the data acquired in a circular cone-beam scan are only sufficient for accurate image reconstructions in the middle of the volume. It is a well-known fact that image reconstruction at the periphery of the volume suffer from cone-beam artifacts.42 Thus, the location of the cervical spine, because it is much off the center of the volume, could have caused the inferior results.

As with every cadaver study, our research had some limitations. One possible constraint is whether the maceration might have influenced the bone properties that would falsify the anatomic reference. Enzymatic maceration, in contrast to common maceration with aggressive solutions, is an established method for removal of soft tissues while maintaining the structural integrity of compact bone43,44 and produces excellent results when applied to ethanol-fixed material.45 Therefore, it seems safe to presume that the assessment of osseous anomalies in enzymatic macerated specimens is admissible. The low number of specimens might also be considered a limitation, since it precludes the possibility of carrying out more comprehensive statistical testing. Engaging more radiologists to assess the sample would certainly generate sufficient data to account for statistical differences in hypothesis testing and would additionally reduce the uncertainty (ie, the standard error) in the descriptive analysis. It has been shown that as the number of raters increases, the required number of subjects decreases. But the savings in sample size obtained by increasing the number of raters reportedly diminishes rapidly after the accrual of 5 raters.46 Mindful of this constraint, we designed this study to produce no more than a descriptive analysis. Nonetheless, our results convey clear answers to the objectives of this study, making further statistical testing or more specimens unnecessary: lateral cephalograms can evidently cause false-positive findings of fusions, and the reported standard errors show that the descriptive values are accurate enough to draw clear conclusions about the MDCT and CBCT evaluations.

CONCLUSIONS

1. Because only a few specimens were evaluated, no general conclusions can be drawn about the prevalence. Yet, lateral cephalograms have been proven to cause false-positive detection of fusions. Lateral cephalograms are therefore a questionable means to assess cervical spine anomalies, and previous studies evaluating fusions in cervical spines, based on 1 lateral cephalogram, seem to be highly problematic.

2. Both MDCT data viewed by general radiologists and CBCT data screened by oral radiologists are reliable methods to exclude fusions.

3. General radiologists appraising MDCT data performed better in the assessment of osteoarthritic changes of the joints than did oral radiologists with CBCT data, but further studies with more specimens would be welcomed to confirm this finding.

We thank Philippe Halioua for the outstanding photographs and Sabrina Beutler and Axel Lang for macerating the specimens.

REFERENCES


