Inter- and Intra-Observer Variability of the Volume of Cervical Ossification of the Posterior Longitudinal Ligament Using Medical Image Processing Software

Dong Ah Shin, M.D., Ph.D.,¹ Gyu Yeul Ji, M.D.,¹,² Chang Hyun Oh, M.D.,² Keung Nyun Kim, M.D.,¹ Do Heum Yoon, M.D.,¹ Hyunchul Shin, M.D., Ph.D.³

Department of Neurosurgery,¹ Yonsei University College of Medicine, Seoul, Korea
Spine and Joint Research Institute,² Guro Cham Teun Teun Hospital, Seoul, Korea
Department of Neurosurgery,³ Kangbuk Samsung Hospital, Sungkyunkwan University College of Medicine, Seoul, Korea

Objective: Computed tomography (CT)-based method of three dimensional (3D) analysis (MIMICS®, Materialise, Leuven, Belgium) is reported as very useful software for evaluation of OPLL, but its reliability and reproducibility are obscure. This study was conducted to evaluate the accuracy of MIMICS® system, and inter- and intra-observer reliability in the measurement of OPLL.

Methods: Three neurosurgeons independently analyzed the randomly selected 10 OPLL cases with medical image processing software (MIMICS®) which create 3D model with Digital Imaging and Communication in Medicine (DICOM) data from CT images after brief explanation was given to examiners before the image construction steps. To assess the reliability of inter- and intra-examiner intraclass correlation coefficient (ICC), 3 examiners measured 4 parameters (volume, length, width, and length) in 10 cases 2 times with 1-week interval.

Results: The inter-examiner ICCs among 3 examiners were 0.996 (95% confidence interval [CI], 0.987–0.999) for volume measurement, 0.973 (95% CI, 0.907–0.978) for thickness, 0.969 (95% CI, 0.895–0.993) for width, and 0.995 (95% CI, 0.983–0.999) for length. The intra-examiner ICCs were 0.994 (range, 0.991–0.996) for volume, 0.996 (range, 0.944–0.998) for length, 0.930 (range, 0.873–0.947) for width, and 0.987 (range, 0.985–0.995) for length.

Conclusion: The medical image processing software (MIMICS®) provided detailed quantification OPLL volume with minimal error of inter- and intra-observer reliability in the measurement of OPLL.

Key Words: Ossification · Posterior longitudinal ligament · Volume · 3D analysis · MIMICS.
INTRODUCTION

Ossification of the posterior longitudinal ligament (OPLL) of the cervical spine is one of the most important causes of cervical myelopathy, and it is more commonly observed in Asian population. The size of OPLL is one of important factor relating to spinal canal stenosis and causing myelopathy. So, evaluating the size and growth of OPLL is crucial to determine the timing of the operation and the risk factors for rapid progression of OPLL. Previous attempts have been made to measure the size of OPLL using the Tsuyama system, but, these methods provided limited information by involving two-dimensional (2D) imaging. The technical improvements in computed tomography (CT)-based three-dimensional (3D) imaging analysis have made accurate 3D measurement of OPLL possible. Recently, a study reported that novel CT-based method of 3D analysis (MIMICS, Materialise, Leuven, Belgium) provided the detailed OPLL classification and quantification in the ossified volume. This method appeared to be very useful for evaluation of OPLL, but accidental errors could be occurred because the identification of ossification was not completely automatic. So, its reliability and reproducibility are obscure. Hence, the present study was designed to evaluate the accuracy of MIMICS® system to evaluate the OPLL, and inter- and intra-observer reliability of this method in the volume of OPLL were measured.

Table 1. Inter-examiner reliability of each measurement

<table>
<thead>
<tr>
<th>Data set</th>
<th>ICC</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volume</td>
<td>0.996</td>
<td>0.987–0.999</td>
</tr>
<tr>
<td>Thickness</td>
<td>0.973</td>
<td>0.907–0.994</td>
</tr>
<tr>
<td>Width</td>
<td>0.969</td>
<td>0.895–0.993</td>
</tr>
<tr>
<td>Length</td>
<td>0.995</td>
<td>0.983–0.999</td>
</tr>
</tbody>
</table>

Table 2. Intra-examiner reliability of each measurement

<table>
<thead>
<tr>
<th>Examiner</th>
<th>Volume</th>
<th>Thickness</th>
<th>Width</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>SHC</td>
<td>0.996</td>
<td>0.998</td>
<td>0.947</td>
<td>0.995</td>
</tr>
<tr>
<td>CPG</td>
<td>0.995</td>
<td>0.944</td>
<td>0.873</td>
<td>0.985</td>
</tr>
<tr>
<td>PSH</td>
<td>0.991</td>
<td>0.994</td>
<td>0.913</td>
<td>0.978</td>
</tr>
<tr>
<td>Range</td>
<td>0.991–0.996</td>
<td>0.944–0.998</td>
<td>0.873–0.947</td>
<td>0.985–0.995</td>
</tr>
</tbody>
</table>

MATERIALS AND METHODS

Study design and subjects enrolment

A total of 10 cases (5 male and 5 female) that underwent surgery for cervical OPLL from 1999 to 2013 was randomly selected from the OPLL database in this analysis. This study was approved by the institutional review board of university hospital and the informed consent was obtained from all patients. Three neurosurgeons participated in this study as

Fig. 1. Steps for 3D reconstruction and measurement. A : CT scanning with 3 mm thickness. B : The obtained images are transmitted digitally via PACS. C : 3D software imports CT data in the DICOM format. D : The segmented object (visualized by a colored-mask) is obtained by including only those pixels of the image with a value higher than or equal to the threshold value (thresholding). E : The segmented object is further split by making another mask on OPLL and removing floating pixels (region growing). F : 3D representation is created and the volume is calculated automatically by the software. CT : computed tomography, DICOM : Digital Imaging and Communications in Medicine, OPLL : ossification of the posterior longitudinal ligament.
examiners. They had no prior experience or training of MIMICS® software which create 3D model with digital imaging and communications in medicine (DICOM) data from CT images. Only the brief explanation was given to these examiners just before the image construction steps.

Fig. 2. Reconstructed 3D images of OPLL after the calculation of the 3D. A: Anterior view. B: Lateral view. C: Posterior view. D: A volumetric mesh with the material assignment for finite element analysis. OPLL: ossification of the posterior longitudinal ligament.

Fig. 3. Measurements. A: The width was defined as the maximum distance along the axis of the coronal plane of OPLL and the thickness was defined as the maximum distance along the axis of the sagittal plane of OPLL. B: The length was defined as the longest dimension of OPLL. C: The volume was calculated automatically by the software in the 3D-reconstructed image. OPLL: ossification of the posterior longitudinal ligament.
Image acquisition and data transfer

CT images of the cervical spine were obtained using a CT scanner (Brivo; GE Healthcare, Waukesha, WI, USA) under uniform conditions (slice thickness, 1.5 mm; voltage, 120 kV; current, 75 mA). The images were saved in the DICOM format and transmitted into a workstation (Z820; HP, Palo Alto, CA, USA) via PACS. Then medical image processing software (MIMICS®) imported the data for 3D reconstruction and measurement. The time was calculated from the step of importing CT data into the software to the end of measurement.

3D reconstruction and measurement

The process consisted of 4 steps (Fig. 1). First, the spine was segmented by threshold which is the function of including only those pixels of the image with a value higher than or equal to the threshold value (segmentation). Second, the segmented object was further separated into the vertebra and the OPLL by making another mask on the OPLL and removing floating pixels (region growing). Whereas initial segmentation is easily achieved by threshold function on a set threshold value, further separation needs diligent work of removing the vertebra attached to the OPLL. The separation was difficult due to obscure boundary between the vertebra and the OPLL. Third, a 3D object of OPLL was reconstructed (3D reconstruction) (Fig. 2). Fourth, the parameters were measured on the processed images (measurement) (Fig. 3). The width was defined as the maximum distance along the axis of the coronal plane of OPLL. The thickness was defined as the maximum distance along the axis of the sagittal plane of OPLL. The length was defined as the longest dimension of OPLL. The measurement speed became quicker from 24.7±4.1 minutes initially to 22.7±4.3 minutes afterwards (p=0.003). The rate-limiting step was the separation step to split the vertebra and OPLL. The mean volume was 1915.2±1027.3 mm$^3$ (range, 780.2–4123.1 mm$^3$). The mean width was 11.0±2.9 mm (range, 5.4–15.9 mm). The mean thickness was 3.6±1.3 mm (range, 2.0–6.3 mm). The mean length was 27.6±10.4 mm (range, 11.2–49.5 mm). There was no significant differences in volume, width, thickness, and length among the examiners (p=0.969; p=0.539; p=0.712; p=0.878).

Validation of the measurement

The ICC among three examiners for the measurement of volume was 0.996 (95% CI, 0.987–0.999). The ICC for thickness was 0.973, and the CI was 0.907–0.978. The ICC for width was 0.969, and the CI was 0.895–0.993. The ICC for length was 0.995, and the CI was 0.983–0.999. The ICC between the two measurement by the same examiner was 0.994 (range, 0.991–0.996) for volume, 0.996 (range, 0.944–0.998) for length, 0.930 (range, 0.873–0.947) for width, and 0.987 (range, 0.985–0.995) for length, indicating high intra-examiner reliability.

RESULTS

Measurement process

The sample consisted of 10 subjects (5 male and 5 female). The mean age was 49.5±10.9 years (range, 36–63 years). The mean time for measuring each case was 23.7±4.3 minutes (range, 15–30 minutes). The measurement speed became quicker from 24.7±4.1 minutes initially to 22.7±4.3 minutes afterwards (p=0.003). A 95% confidence level was used for all tests. All data were analyzed by SPSS ver. 18 (SPSS Inc., Chicago, IL, USA). Values are reported as the mean±standard deviation. Each of statistical method is described at the end of each result. A p-value of <0.05 was considered statistically significant.

Validation of the measurement

To assess the reliability of the measurement, 3 examiners measured 4 parameters in 10 cases 2 times with 1 week interval. Inter-examiner and intra-examiner intraclass correlation coefficient (ICC) and 95% CI were assessed. Inter-examiner reliability was assessed by the ICC of data obtained for the measurement of each parameter among all three examiners.

Intra-examiner reliability was also assessed between the first and second measurements by all three examiners.

DISCUSSION

Measurements of ossification length and thickness using 2D...
X-ray images are possible to a certain extent\(^1,3,5,21\), but, 2D evaluation with plain lateral radiographs has some limitations. First, its accuracy is reduced due to errors associated with the imaging procedure, because the ossification width and volume are not measurable. Ossification not identifiable by X-ray examination can be evaluated on CT images, and therefore, the classification discrimination may differ from that obtained by X-ray imaging. Second, the growth can occur in any direction, only cranio-caudal and ventrodorsal progression are depicted on plain radiographs. Growth in the oblique direction is projected only orthographically onto a radiograph. Indeed, multidimensional evaluation of ossification can be achieved with CT images, but it is difficult to evaluate continuity on 2D X-ray images. Other disadvantages of X-ray analysis include inability to identify small-area ossifications and to accurately locate the ossification.

Compared with radiography, CT scanning has several advantages. CT scanning is the most sensitive diagnostic method for detecting small ossifications or calcifications of the ligament, which are likely to be missed on radiographs\(^19\). However, conventional CT scanning also has some limitation of slice thickness. Measurement of OPLL on conventional radiographs is unreliable because the slice is likely to be thick, owing to the limited number of films used. It is also difficult to slice OPLL on the same level and angle to compare past and present ossification. However, these limitations can be overcome using helical scanning with multi-detector CT and digital viewers, but, its reliability and reproducibility are obscure\(^2\).

Three-dimensional evaluation with CT scanning is better suited to measuring OPLL growth. OPLL could be masked by the shoulder girdle shadows in the lower cervical spine (Fig. 4), and might be less distinct because the ossification is less densely calcified in young patients\(^7\). CT based method of 3D analysis can precisely evaluate the volume of OPLL at one point and provide information about the volume of OPLL with minimal analytical error. Evaluation of continuity and classification of ossification are comparatively simple in 3D images\(^2\). With CT images, changes in ossification form and thickness on the caudal side can be evaluated in detail in patients with ossifications growing toward the cranial side. Slight graininess of the images associated with slice thickness is seen at ossification boundaries. Hence, the accurate superimposition by voxel based registration could provide comparison of OPLL before and after surgery. Indeed, it is also possible to derive an accurate numerical estimate of both overall absolute value of the volume of ossification and its rate of progression over time, and to compare these volumes with those derived from subsequent scans by calculating exact volumes.

Izumi et al.\(^11\) previously reported CT-based 3D analysis method to measure volume changes in OPLL. The subjects were 20 OPLL patients who were being followed using the MIMICS\(^\text{®}\) software (slice thickness, 1.25 mm; voltage, 120 kV; current, 178 mA or slice thickness, 0.5–1.0 mm; voltage, 120 kV; current, 75 mA) to calculate the volume. They concluded that this 3D method used allowed detailed OPLL classification and quantification of change in the ossified volume and ap-
appears to be very useful for quantitative evaluation of OPLL with only minimal measurement error. The ICC calculated in this study was obviously high, suggesting that the evaluation of ossification volume was accurate and valid, but, they also commented that accidental errors might have occurred because this method used for identification of ossification was not completely automatic.

In the study to evaluate the inter- and intra-observer reliability of lateral radiograph, axial CT, 2D and 3D reconstructed CT images based OPLL classification, the inter- and intra-observer ICCs were only 0.51 and 0.67 for the lateral radiograph, even in combination with the axial CT images, 0.70 and 0.85 for 2D CT images, and 0.76 and 0.86 for 3D CT images, respectively. These values showed a good to excellent range for the 2D and 3D reconstructed CT images while those of the lateral radiograph indicated a fair range. These ICCs could improve the correlation by using software system (Visualization Toolkit (http://www.vtk.org/)) as 0.968 (95% CI, 0.880–0.992) for inter-observer reproducibility and 0.987 (95% CI, 0.953–0.997) for intra-observer reproducibility of volume measurement. Recent CT-based method of 3D analysis using MIMICS® software also provided the detailed OPLL quantification in the ossified volume with excellent inter- (>0.969) and intra-observer reproducibility (>0.930) although the examiners were not professional users. Similarly result was also reported in the study using MIMICS® with 0.856 of inter-observer ICC and 0.999 of intra-observer ICC. Therefore, this measurement method seems to be universally applicable method using similar viewers with equivalent function.

CONCLUSION

The medical image processing software (MIMICS®) provided detailed quantification of OPLL volume with minimal error of inter- and intra-observer reliability in the measurement of OPLL.

References


