

RESEARCH ARTICLE

Quantitative assessment of condylar bone resorption using fused CBCT images: differentiating and diagnosing three distinct groups based on volume and thickness decrease

¹Ji-ling Feng, ¹Ruo-han Ma, ¹Li-li Sun, ¹Jun-ru Zhao, ^{1,2}Yan-ping Zhao and ¹Gang Li

¹Department of Oral and Maxillofacial Radiology, Peking University School and Hospital of Stomatology & National Center for Stomatology & National Clinical Research Center for Oral Diseases & National Engineering Research Center of Oral Biomaterials and Digital Medical Device & Beijing Key Laboratory of Digital Stomatology & Research Center of Engineering and Technology for Computerized Dentistry Ministry of Health & NMPA Key Laboratory for Dental Materials, Beijing, China; ²Center for Temporomandibular Disorders and Orofacial Pain, Peking University School and Hospital of Stomatology, Beijing, China

Objectives To investigate the accuracy of fused CBCT images in diagnosing three distinct groups of bone changes characterized by volume and thickness decrease in patients with temporomandibular joint osteoarthritis (TMJ OA) during follow-up.

Methods In this retrospective study, 109 patients (176 TMJs) with TMJ OA were included. Two consecutive CBCT images for the same patient were registered and fused. Then, three image sets were established: without fusion, fused 2D image, and fused 3D image. Three residents randomly and independently evaluated whether there was condylar resorption with the three image sets respectively. The samples diagnosed as condylar resorption by the expert panel were divided into three subgroups according to the volume and thickness decrease calculated after segmentation. The inter- and intraobserver agreement, receiver operating characteristic (ROC), and area under the curve (AUC) evaluated the diagnostic capability for different subgroups.

Results For the volume decrease more than 50 mm³ and thickness decrease more than 1 mm groups, the AUC values for fused image sets were higher than those without fusion ($p < 0.01$). For the volume decrease within 50 mm³ and thickness decrease within 1 mm groups, the AUC values for fused 2D image sets were higher than the image sets without fusion ($p < 0.05$), but there was no significant difference between the fused 3D image sets and the image sets without fusion ($p = 0.48$ for volume decrease, $p = 0.37$ for thickness decrease).

Conclusions The fused images can improve the diagnostic accuracy and repeatability for the samples with at least 50 mm³ volume decrease or 1 mm thickness decrease compared with the image groups without fusion.

Dentomaxillofacial Radiology (2023) 52, 20230337. doi: [10.1259/dmfr.20230337](https://doi.org/10.1259/dmfr.20230337)

Cite this article as: Feng J, Ma R, Sun L, Zhao J, Zhao Y, Li G. Quantitative assessment of condylar bone resorption using fused CBCT images: differentiating and diagnosing three distinct groups based on volume and thickness decrease. *Dentomaxillofac Radiol* (2023) 10.1259/dmfr.20230337.

Keywords: Cone-Beam Computed Tomography; Temporomandibular joint; Radiostereometric analysis; Diagnostic Imaging

Introduction

As the growth and development center of mandible,¹ condyle modifies its shape throughout an individual's entire lifetime.² Some evidence suggests that excessive overloads, which exceed physiological capacity, can lead to degenerative remodeling of the temporomandibular joint (TMJ).³⁻⁵ Temporomandibular joint osteoarthritis (TMJ OA) is a degenerative remodeling type affecting approximately 15.9% of adults and elderly individuals and 0.2% of children and adolescents.⁶ The main changes experienced by patients with TMJ OA, such as erosions, osteophytes, flattening, sclerosis, and cyst-like change,^{7,8} occur in bone. CBCT with a higher spatial resolution and lower radiation dose performs well in diagnosing bone change when compared with helical CT.⁹⁻¹²

TMJ OA is a chronic and long-lasting disease. Follow-up visits for this condition often involve a repeat CBCT image examination to check for any changes in the condylar bone. Currently, the common clinical comparison method is to directly observe the two CBCT images. Researchers have conducted various studies¹³⁻¹⁹ comparing condylar bone change during follow-up, which can be categorized into two groups: direct observation^{13,14} and observation using registration methods.¹⁵⁻¹⁹ Direct observation may not be sufficient when the bone change is minor, or the condylar edge is irregular. In such cases, methods of observation after registration can be used. Schilling *et al*²⁰ reported that this method is reliable and can be used to quantify even subtle bony differences in the three-dimensional (3D) condylar morphology. Feng *et al*¹⁹ further reported that fused two-dimensional (2D) cross-sectional and 3D images can provide higher accuracy and consistency in diagnosing condylar bone resorption compared with direct observation during follow-up. Although using fused image for diagnosing condylar bone resorption can achieve high diagnostic accuracy and consistency, there is still a lack of quantitative evidence to determine how much the difference in bone volume decrease can be differentiated and still maintain that accuracy and consistency after image fusion.

Thus, the main purpose of the present study was to explore how accurately a small range of volume decrease of the condylar bone during the follow-up can be detected. To achieve this purpose, we divided the samples with bone resorption into three groups based on volume and thickness decrease, respectively.

Methods and materials

Subjects

According to the previous study,¹⁹ the prevalence of bone resorption was 57%. Assuming that the proportional incidence of bone decrease more than 100mm³, between 50 and 100mm³, and less than 50mm³ is 19%

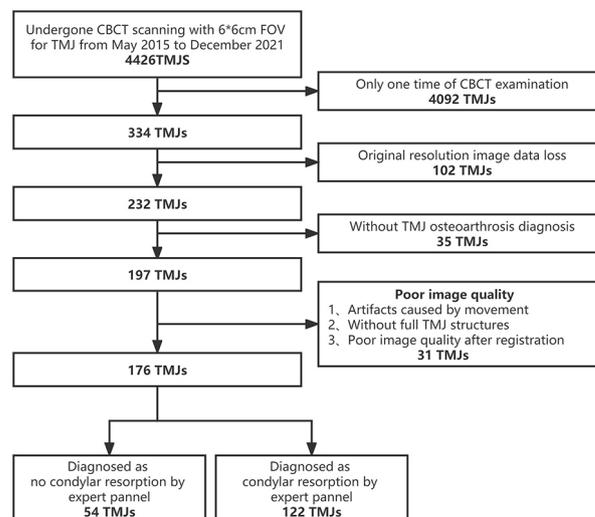


Figure 1 Flowchart of study subject selection. TMJ, temporomandibular joint.

respectively, and the proportional incidence of the three groups with thickness decrease of more than 2mm, between 1 and 2mm and less than 1mm is also 19% respectively, at least 17 positive cases for each group were needed when a sensitivity and specificity of 0.85 for fusion images is considered for the detection of bone decrease in different groups.

In the present study, we collected a total of 122 positive cases and 54 negative cases as shown in [Figure 1](#).

This study was registered for the WHO international clinical trial (ChiCTR2200060049), filed in the Chinese Medical Research Registration and Filing System (MR-11-23-016454), and approved by the Institutional Review Board of Peking University School and Hospital of Stomatology (PKUSSIRB- 201944056). This study included 176 TMJs from 109 patients who visited the Center for Temporomandibular Disorders and Orofacial Pain at Peking University School and Hospital of Stomatology from 2015 to 2021. The other 42 condyles of 109 subjects were not included due to the absence of imaging changes in bone on CBCT images at both time points. The inclusion criteria were as follows: (1) the patients have been diagnosed with TMJ OA according to Diagnostic Criteria for Temporomandibular Disorders (DC/TMD)²¹; (2) the acquired CBCT image data included the full TMJ structures; (3) the patients' age was more than 25 years. The exclusion criteria were as follows: (1) metal or motion artifacts influence the quality of CBCT image data; (2) the interval of two CBCT examinations was less than 3 months. The 3-month cut-off was chosen for one reason that the radiographic manifestations of bone resorption often require a certain amount of time to develop, and the other is that a 3-month interval aligns with the clinical follow-up intervals.

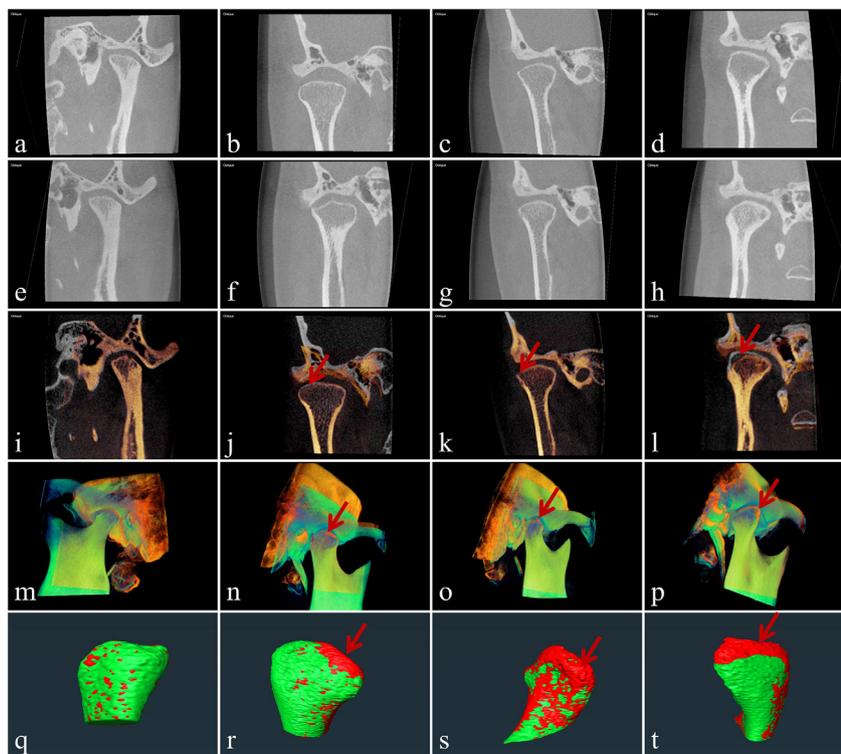


Figure 2 Example images of the condyles with different ranges of volume decrease. The oblique coronal images were scanned at the first visit (a–d), at the second visit (e–h), the fused 2D cross-sectional images (i–l), the fused 3D images (m–p), and the fused 3D images after the segmentation of condyles. a, e, i, m, and q were samples without bone change; b, f, j, n, and r were samples with volume decrease within 50 mm^3 ; c, g, k, o, and s were samples with volume decrease between 50 and 100 mm^3 ; d, h, l, p, and t were samples with volume decrease more than 100 mm^3 . The red arrow indicates bone resorption. 2D, two-dimensional; 3D, three-dimensional.

CBCT image acquisition

The patients kept maximal intercuspation in a sitting position with the Frankfort plane paralleling to the floor during the CBCT scan. The thyroid collar was worn to reduce the radiation to patient. The CBCT unit used was 3D Accuitomo 170 (J Morita Mfg., Corp., Kyoto, Japan). The exposure parameters were as follows: scanning time of 17.5s, tube current of 5–6 mA, tube potential of 85–90 kVp, and field of view (FOV) was $6 \times 6\text{ cm}$, which included one side of TMJs and part of the mandible.

All the CBCT image data sets were reconstructed with a voxel size of 0.125 mm and exported in Digital and Communication in Medicine (DICOM) format.

Image registration

The registration of two consecutive CBCT image data sets from the same patients was carried out with the Amira visual software (v. 2020.2, Thermo Fisher Scientific, France) in the multiplanar viewer module. The registration included three steps: manual registration, auto registration, and manual adjustment. The rigid transform model was adopted for this registration, and the voxel registration was performed based on the mandibular ramus and the coronoid process, which were relatively stable parts for patients with TMJ OA.

The first CBCT image data set was set as the primary data and the second data set was set as the overlay data. More detailed information could be found in the previous study.^{19,22}

To evaluate the registration accuracy, two oral and maxillofacial radiologists with 4 and 10 years of experience in CBCT image interpretation evaluated the registered images using the subjective evaluation method of a 5-point scale (1-very poor, 2-poor, 3-average, 4-good, 5-very good) randomly and independently. All the samples were kept for further analysis when the evaluation points reached 4 or 5. As for the samples with scores of 1, 2, 3, a manual adjustment was used to additionally adjust the registration. In case that the evaluation point does not reach score 4 and 5 neither, the images were not used for further assessment.

To evaluate the registration repeatability, 104 data sets from 52 TMJs were randomly selected to register one more time after 2 weeks. The displacement translation values in the X, Y, and Z axis, and the position and rotation angle of the X, Y, and Z rotation axis of 52 overlay data sets were recorded.

Reference standard

The reference standard was acquired by the expert panel. Two oral and maxillofacial radiologists with 10

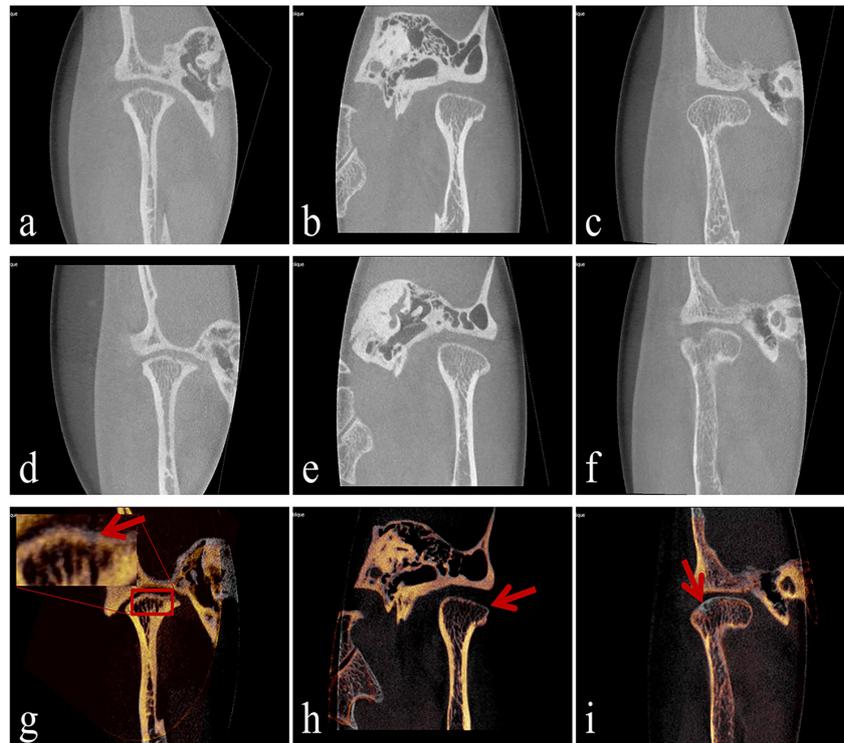


Figure 3 Example images of the condyles with different ranges of thickness decrease. The oblique coronal images were scanned at the first visit (a–c), at the second visit (d–f), and the fused 2D cross-sectional images (g–i). a, d, and g were samples with a thickness decrease within 1 mm; b, e, and h were samples with a thickness decrease between 1 and 2 mm; c, f, and I were samples with a thickness decrease of more than 2 mm. The red arrow indicates bone resorption. 2D, two-dimensional.

and 25 years of experience each, and one expert with an experience of over 30 years in oral and maxillofacial surgery (mainly for the diagnosis and conservative treatment of TMJ disease) acted as the expert panel. Two options, yes or no, were presented to two experts to independently determine whether any condylar resorption had occurred for the second CBCT images according to the two consecutive CBCT images without fusion, fused 2D cross-sectional, and fused 3D images. In case that the determinations were not the same, the specialist with an experience of over 30 years in oral and maxillofacial surgery was involved in negotiation and confirmation.

Quantitative measurement based on volume decrease

The segmentation process was conducted with segmentation module in the Amira visual software. Before segmentation, the overlay CBCT data set was resampled with respect to the primary CBCT data set after registration. The segmentation includes two steps: automatic segmentation and manual correction. For the automatic segmentation, the lower and upper gray level threshold was set up first. The lower threshold was between 200 and 400 and the upper threshold was set as 1500. For the same TMJ scanned at two different times, the lower gray level threshold was set at the same threshold value to reduce the possible effect on condyle volume calculations. After automatic segmentation, the investigator manually modify the segmentation labels layer by layer

in three planes of space (axial, oblique sagittal, oblique coronal). The lower boundary of the segmentation was the sigmoid notch. Before the segmentation, the CBCT data obtained from the second time point were aligned and resampled with the CBCT data from the first time point, which was considered as the basis. Thus, when the lowest layer of the Z-axis of the image data segmented at the two time points was the same, the two consecutive CBCT image was segmented at the exactly the same lower boundary (Figure 2q–t).

After the segmentation, the volume of the condyle above the sigmoid notch was calculated. Then, the volume difference of the same condyle scanned at two times was calculated. All the samples with bone resorption determined by the expert panel were divided into three groups: volume decrease more than 100 mm³ (Figure 2d, h, l, p and t); volume decrease between 50 and 100 mm³ (Figure 2c, g, k, o and s); volume decrease less than 50 mm³ (Figure 2b, f, j, n and r).

All the segmentation process was completed by the same investigator. To verify the repeatability of the segmentation, 52 data sets from 26 TMJs were randomly selected to segment one more time after 2 weeks. The volumes were recorded.

Quantitative measurement based on thickness decrease

The thickness decrease for the samples with bone resorption was measured in the coronal view of the

Table 1 Demographic feature, volume and thickness decrease characteristics of the study sample

	<i>Total samples</i>	<i>Positive samples</i>	<i>Negative samples</i>	<i>p-value</i>
Number of patients	176	122	54	NA
F/M ratio	163:13	118:4	45:9	0.003
Age \pm SD	36 \pm 10	35 \pm 10	37 \pm 12	0.051
Interval time (month)				
Average \pm SD	15.7 \pm 12.9	16.7 \pm 13.9	13.5 \pm 10.1	0.092
Median (IQR)	10.5 (6.0–23.0)	11.0 (6.0–27.0)	9.0 (5.0–22.3)	
Average VD (mm ³)				
Average \pm SD	64.63 \pm 109.96	95.35 \pm 118.26	-4.78 \pm 30.24	0.000
Median (IQR)	29.50 (-6.75–115.5)	70.00 (6.75–144)	-5.00 (-34.50–24.75)	
Average TD (mm)				
Average \pm SD	1.22 \pm 1.36	1.76 \pm 1.31	0 \pm 0	0.000
Median (IQR)	0.88 (0.00–1.85)	1.34 (0.80–2.55)	0 (0–0)	

IQR, interquartile range; NA, not applicable; Negative samples, samples without condylar bone decrease; Positive samples, samples with condylar bone decrease; SD, standard deviation; TD, thickness decrease; VD, volume decrease.

multiplanar viewers after registration. The first step was to adjust the 2D threshold as 0–1500, the second step was to correct the coronal image to the largest observation plane, and the last step was to record the 3D coordinates of the two farthest points reflecting thickness changes in the resorption area of the condylar bone. Then, the distance was calculated using a formula. All measurement process was repeated three times by the same investigator to reach an average value.

All the samples with bone resorption determined by the expert panel were divided into three groups: thickness decrease more than 2 mm (Figure 3c, f and i; thickness decrease between 1 and 2 mm (Figure 3b, e and h) and thickness decrease less than 1 mm (Figure 3a, d and g).

Image evaluation

Three oral and maxillofacial radiology residents with different years of experience (1, 4, and 4 years respectively) acted as observers. All the observers were calibrated with an additional session of images before evaluation. The theory of image registration and fusion, the use of software, and evaluation methods were explained and demonstrated before the formal experiment. Three sets of images ((1) CBCT images at two time points without fusion; (2) fused 2D cross-sectional CBCT images; (3) fused 3D CBCT images) for 176 TMJs were evaluated randomly and independently. There was at least 1 week apart between any two sets of image evaluations. There were five choices offered to the

observers for the evaluation of condylar resorption: (1) definitely no resorption; (2) probably no resorption; (3) questionable; (4) probably resorption; and (5) definitely resorption.

While evaluating, the observers could adjust the images' position, angle, and magnification freely with no time limits in a quiet and dim room. The monitor used during the observation was Nio Color 5.8 MP (MDNC-6121) display (Barco NV, Kortrijk, Belgium).

For the evaluation of intraobserver consistency, 89 samples were randomly stratified from 176 samples for the secondary evaluation, which include 27 samples without condylar resorption, 25 samples with volume decrease more than 100 mm³, 10 samples with volume decrease between 50 and 100 mm³, and 27 samples with volume decrease within 50 mm³. The same observers evaluated these samples 4 weeks later under the same condition.

Statistical analysis

Power Analysis and Sample Size (PASS) software package v. 21.0.3.0 (NCSS, LLC, Kaysville, UT) was used for sample size calculation. IBM SPSS Statistics for Windows v. 25.0 (IBM Corp., Armonk, NY) and MedCalc Statistical Software v. 20.027 (MedCalc Software bvba, Ostend, Belgium) were used for data analysis.

Wilcoxon signed-rank test was used to determine the statistical significance of the repeatability of registration. A *p*-value of 0.05 or less was considered significant.

Table 2 Wilcoxon signed ranks test for the repeatability of registration

	<i>TX2 - TX1</i>	<i>TY2 - TY1</i>	<i>TZ2 - TZ1</i>	<i>RX2 - RX1</i>	<i>RY2 - RY1</i>	<i>RZ2 - RZ1</i>	<i>R2 - R1</i>
Z	-.929 ^a	-.920 ^a	-1.402 ^b	-.993 ^b	-1.038 ^a	-.410 ^b	-.756 ^b
Asymp.Sig. (2-tailed)	.353	.358	.161	.321	.299	.682	.450

R1, R2, the rotation degree for the first and second registration; RX1, RY1, RZ1, RX2, RY2, RZ2, the rotation axis of X, Y, Z-axis for the first and second registration; TX1, TY1, TZ1, TX2, TY2, TZ2, the translation of X, Y, Z-axis for the first and second registration.

^aBased on negative ranks.

^bBased on positive ranks.

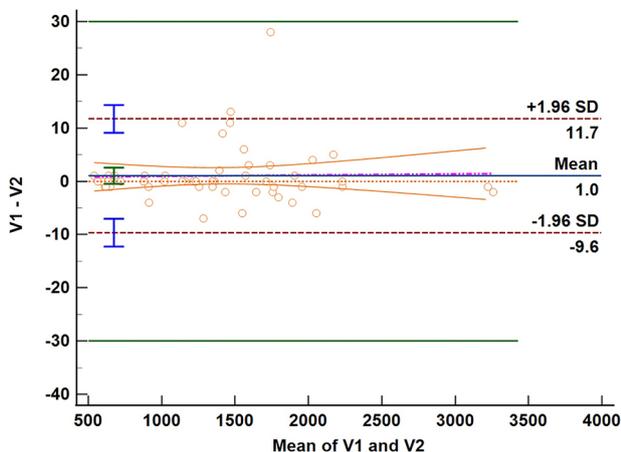


Figure 4 The Bland–Altman plot for the volume difference of repeated segmentation. The green lines are the maximum allowed difference between two times of segmentation, the red dashed lines represent the lower and upper limit of agreement, and the blue line represents the average difference. The orange lines were the 95% CI of the regression line of difference. V1, V2 (mm³): the volume calculated by the first and second segmentation. CI, confidence interval.

The Bland–Altman analysis was used to verify the repeatability of segmentation results.

The intra- and interobserver agreement was assessed by intraclass correlation coefficients (ICCs). ICC values and their 95% confidence intervals were calculated based on a single-rating (intra-agreement)/mean rating (inter-agreement), absolute-agreement, and 2-way mixed-effects model. The ICC values were interpreted as poor (<0.50), moderate (0.50–0.75), good (0.75–0.90), or excellent (>0.90) in agreement.²³

The receiver operating characteristic curves (ROCs) was performed with MedCalc Statistical Software for the representation of diagnostic accuracy of the six image sets.

Results

The demographic and volume/thickness decrease characteristics is shown in Table 1. Totally, 176 samples were collected, which included 122 positive samples and 54 negative samples. The mean age was 36 ± 10 years and 163 samples were from females and 13 samples were from males. The average interval time between the

two CBCT scans of the same patients was 15.7 ± 12.9 months.

Results from the Wilcoxon signed ranks test showed no significant difference between the two registrations (Table 2).

The Bland–Altman analysis indicates the volume differences from the two segmentations. 50 out of 52 points were within the 95% limits of agreement (Figure 4). This result indicates that the segmentation process has a good repeatability (*p* = 0.18)

Interobserver agreement of the three observers is shown in Table 3. All the fused image sets reached good or excellent agreement (ICC ≥ 0.82) for the six groups. For the image set without fusion, the interobserver agreement was good (ICC = 0.79) to the group of volume decrease more than 100mm³, moderate (ICC = 0.58) to the group of volume decrease between 50 and 100mm³, and poor or moderate (ICC = 0.53 (0.32–0.68)) for the group of volume decrease within 50mm³. The same tendency was found for the three groups with different thickness decrease.

Intraobserver agreement of the three observers is shown in Figure 5. For the fused image sets, the intraobserver agreement was good or excellent for the following groups: volume decrease more than 100mm³, between 50 and 100mm³, thickness decrease more than 2mm, between 1 and 2mm. The intraobserver agreement was moderate or good for the volume decrease less than 50mm³ and thickness decrease less than 1mm groups. For the image sets without fusion, the intraobserver agreement was poor or moderate for the six groups with only one exception.

The areas under the ROC curve (AUC) from the pooled data of the three observers are shown in Table 4 and the *p* values when comparing AUC of each image set for the six groups are shown in Table 5. For the volume decrease more than 100mm³, between 50 and 100mm³, thickness decreases more than 2mm, and between 1 and 2mm subgroups, the AUC values for the fused image sets were significantly higher than the image sets without fusion. For the volume decrease within 50mm³ and thickness decrease within 1mm subgroups, there were no significant differences between the fused 3D image sets and the image sets without fusion, but there were still significant differences between the fused 2D cross-sectional image sets and the image sets without fusion. The ROCs from the pooled data of three observers for six groups are shown in Figure 6.

Table 3 Interobserver agreement for each group of the three image sets (ICC [95% CI])

	VD >100mm ³	VD 50–100mm ³	VD <50mm ³	TD >2mm	TD 1–2mm	TD <1mm
Without fusion	0.79 (0.68–0.86)	0.58 (0.35–0.73)	0.53 (0.32–0.68)	0.79 (0.67–0.87)	0.67 (0.50–0.78)	0.45 (0.21–0.62)
Fused 2D	0.95 (0.92–0.96)	0.90 (0.85–0.93)	0.84 (0.78–0.89)	0.94 (0.91–0.96)	0.92 (0.88–0.94)	0.82 (0.75–0.88)
Fused 3D	0.91 (0.86–0.94)	0.86 (0.76–0.91)	0.84 (0.75–0.89)	0.90 (0.84–0.93)	0.88 (0.81–0.92)	0.83 (0.73–0.89)

2D, two-dimensional; 3D, three-dimensional; TD, thickness decrease; VD, Volume decrease.

ICC values, agreement was rated as “poor” (<0.50), “moderate” (0.50–0.75), “good” (0.75–0.90), and “excellent” (>0.90). Without fusion: two times CBCT without fusing; fused 2D, fused image show in 2D cross-sectional; fused 3D: fused image show in 3D.

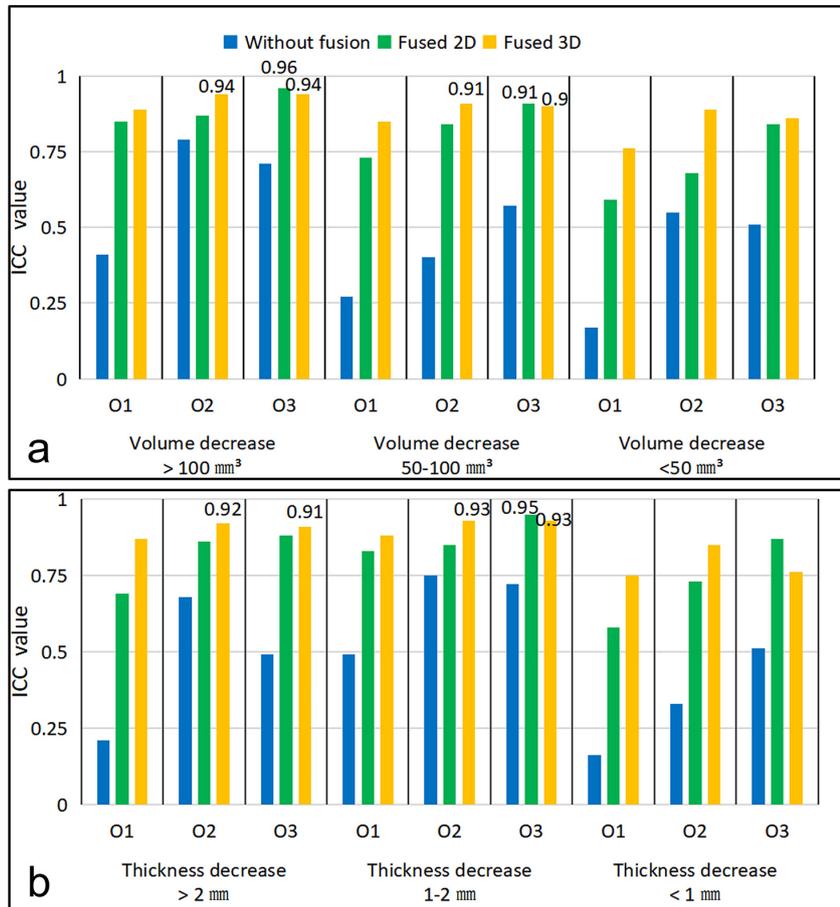


Figure 5 The histogram of the interobserver agreement for six subgroups. (a) The interobserver agreement of three observers in three volume subgroups on the three sets of images. (b) The interobserver agreement of three observers in the three decreased thickness subgroups. O1, observer 1; O2, observer 2; O3, observer 3.

Discussion

In the present study, when the condyle volume decreased by more than 50 mm³ or the thickness decreased by more than 1 mm, the diagnostic accuracy of the fused image sets was significantly higher than that of the image sets without fusion according to AUC values (Table 4, Figure 6a, b, d and e). When the volume decreased within 50 mm³ or the thickness decreased within 1 mm, the diagnostic accuracy was lower than those with volume decrease more than 50 mm³ or the thickness decrease more than 1 mm for the fused

image sets (Figure 6c and f). For all the six groups, the interobserver agreement reached good or excellent for the fused image sets. When the volume decrease was within 100 mm³ or the thickness decreased within 2 mm, the interobserver agreement was poor or moderate for the image sets without fusion. The same tendency was for the intraobserver variances. This indicates that the observers have a good diagnostic accuracy and repeatability in the evaluation of condylar resorption from the fused CBCT images.

Some studies have quantitatively analyzed condylar resorption. Paniagua et al used SPHARM-PDM

Table 4 Area under the receiver operating characteristic curve (AUC) for six groups for the three image sets

	VD > 100 mm ³	VD 50–100 mm ³	VD < 50 mm ³	TD > 2 mm	TD 1–2 mm	TD < 1 mm
Without fusion	0.88 (0.84–0.92)	0.82 (0.76–0.86)	0.69 (0.63–0.74)	0.91 (0.87–0.94)	0.80 (0.74–0.84)	0.67 (0.61–0.72)
Fused 2D	0.97 (0.94–0.99)	0.96 (0.92–0.98)	0.79 (0.75–0.84)	0.97 (0.94–0.99)	0.95 (0.9–0.97)	0.76 (0.71–0.81)
Fused 3D	0.94 (0.91–0.96)	0.93 (0.89–0.96)	0.71 (0.66–0.76)	0.94 (0.91–0.97)	0.89 (0.85–0.92)	0.70 (0.65–0.75)

Fused 2D, fused image show in 2D cross-sectional; Fused 3D, fused image show in 3D; TD, Thickness decrease; VD, Volume decrease; Without fusion, two times CBCT without fusing.

Table 5 values when comparing AUC of each image set for six groups

	VD >100mm ³		VD 50–100mm ³		VD <50mm ³		TD >2mm		TD 1–2mm		TD <1mm	
	Fused 2D	Fused 3D	Fused 2D	Fused 3D	Fused 2D	Fused 3D	Fused 2D	Fused 3D	Fused 2D	Fused 3D	Fused 2D	Fused 3D
Without fusion	<0.05	<0.05	<0.05	<0.05	<0.05	0.48	<0.05	0.13	<0.05	<0.05	<0.05	0.37
Fused 2D		0.049		0.19		<0.05		0.056		<0.05		0.045

Fused 2D, fused image show in 2D cross-sectional; Fused 3D, fused image show in 3D; TD, thickness decrease; VD, volume decrease; Without fusion, two times CBCT without fusing.

(Spherical Harmonics - Point Distribution Model,²⁴ a unique and symmetric point correspondence across all measured surfaces) to quantify temporomandibular joint osteoarthritis and the condyle defects simulated in a thickness of 3 mm and 6 mm, which was greater than the bone reduction thickness explored in this study.²⁵ Schilling et al reported that condylar registration for longitudinal assessments was reliable and the mean interobserver differences were less than 0.6 mm,²⁰ however, further exploration on diagnostic accuracy was lack. Lee et al and Jiang et al applied the color-coded map to qualitatively and quantitatively evaluate condylar bone remodeling after orthognathic/orthodontic treatment.^{17,18} Lee et al reported that the mean of the average point-to-point distances on condylar surface was 0.11 ± 0.03 mm. However, the distance displayed on the color-coded map was not calculated based on two corresponding points, but rather based on the two farthest points that may lead to error. In the present study, the thickness decrease of the condyle was measured using clinical data rather than simulating defects, and the diagnostic accuracy and repeatability were further investigated. These may make the present study more convincing and clinically relevant.

In this study, the minimal volume decrease group was determined as 50 mm³. The reasons were as follows. Firstly, there is an error in the segmentation itself. In the cases diagnosed by the expert panel as having no condylar resorption, the volumes calculated after condylar segmentation at the two time points were mostly different. This difference was maximally within 50 mm³. Secondly, a difference of -10 (-71–50) mm³ between condylar volume which was determined with dry skulls and based on CBCT images has been reported.²⁶

Although the three observers were with only 1 and 4 year experience in oral and maxillofacial radiology, the diagnostic accuracy and intraobserver agreement for the fused image sets were still high. The AUC values were 0.93–0.97 for the volume decrease more than 50 mm³, and was 0.89–0.97 for the thickness decrease more than 1 mm groups. The ICC values for the volume decrease more than 50 mm³ groups were 0.73–0.96 and for the thickness decrease more than 1 mm groups were 0.69–0.95. Therefore, the fused images can help junior residents to accurately diagnose condylar bone resorption.

One limitation of the present study is that the gold-standard cannot be obtained. To solve this problem, the

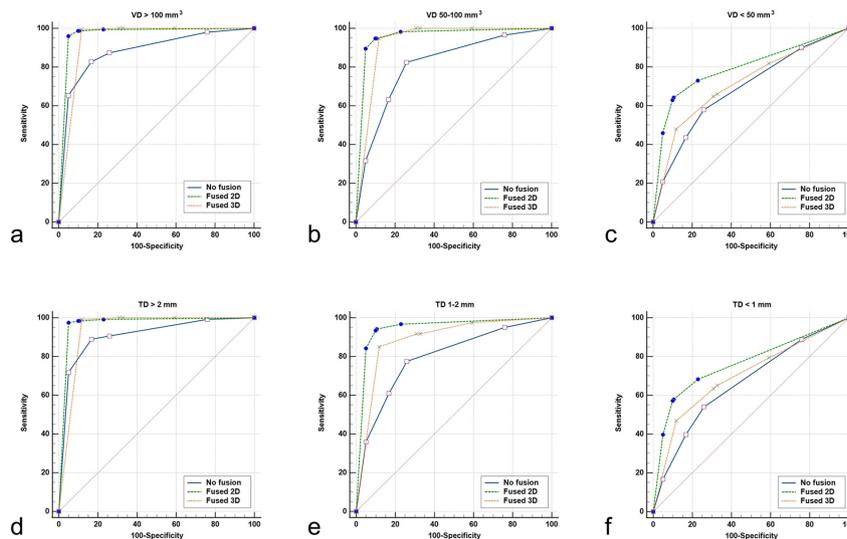


Figure 6 The ROC curves for the three image sets from pooled data of three observers. 2D, two-dimensional; 3D, three-dimensional; ROC, receiver operating characteristic; TD, thickness decrease; VD, volume decrease; a, b, and c are the ROC curves for the three subgroups with an order of volume decrease; d, e, and f are the ROC curves for the three subgroups with an order of thickness decrease.

reference criteria were obtained from a panel of three experts. This is consistent with the approach adopted by some researchers in similar scenarios.^{19,27,28}

In the future, it's necessary to use prospective and multicenter research to validate our methods and improve their applicability in clinical practice. The semi-automatic registration method used in this study takes approximately 10 min to complete the registration of a pair of condyles. The registration steps should also be improved to shorten registration time and ensure registration quality.

Conclusion

The fused 2D and 3D CBCT images can improve the diagnostic accuracy and repeatability for the samples with at least a 50 mm³ volume decrease or a 1 mm thickness decrease compared with the image sets without fusion. This can assist in quantitative evaluation of the disease progression and treatment effectiveness of patients with TMJ OA.

REFERENCES

1. Hinton RJ, Jing Y, Jing J, Feng JQ. Roles of Chondrocytes in Endochondral bone formation and fracture repair. *J Dent Res* 2017; **96**: 23–30. <https://doi.org/10.1177/0022034516668321>
2. Moffett BC, Johnson LC, McCabe JB, Askew HC. Articular remodeling in the adult human Temporomandibular joint. *Am J Anat* 1964; **115**: 119–41. <https://doi.org/10.1002/aja.1001150108>
3. Wang XD, Kou XX, Mao JJ, Gan YH, Zhou YH. Sustained inflammation induces degeneration of the Temporomandibular joint. *J Dent Res* 2012; **91**: 499–505. <https://doi.org/10.1177/0022034512441946>
4. Milam SB, Zardeneta G, Schmitz JP. Oxidative stress and degenerative Temporomandibular joint disease: a proposed hypothesis. *J Oral Maxillofac Surg* 1998; **56**: 214–23. [https://doi.org/10.1016/s0278-2391\(98\)90872-2](https://doi.org/10.1016/s0278-2391(98)90872-2)
5. Mao JJ, Rahemtulla F, Scott PG. Proteoglycan expression in the rat Temporomandibular joint in response to unilateral bite raise. *J Dent Res* 1998; **77**: 1520–28. <https://doi.org/10.1177/00220345980770070701>
6. Valesan LF, Da-Cas CD, Réus JC, Denardin ACS, Garanhani RR, Bonotto D, et al. Prevalence of Temporomandibular joint disorders: a systematic review and meta-analysis. *Clin Oral Invest* 2021; **25**: 441–53. <https://doi.org/10.1007/s00784-020-03710-w>
7. Al-Ani Z. Temporomandibular joint Osteoarthritis: a review of clinical aspects and management. *Prim Dent J* 2021; **10**: 132–40. <https://doi.org/10.1177/2050168420980977>
8. Ondrouch AS. Cyst formation in osteoarthritis. *J Bone Joint Surg Br* 1963; **45**: 755–60.
9. Hintze H, Wiese M, Wenzel A. Cone beam CT and conventional tomography for the detection of morphological Temporomandibular joint changes. *Dentomaxillofac Radiol* 2007; **36**: 192–97. <https://doi.org/10.1259/dmfr/25523853>
10. Honda K, Larheim TA, Maruhashi K, Matsumoto K, Iwai K. Osseous abnormalities of the Mandibular Condyle: diagnostic reliability of cone beam computed tomography compared with Helical computed tomography based on an autopsy material. *Dentomaxillofac Radiol* 2006; **35**: 152–57. <https://doi.org/10.1259/dmfr/15831361>
11. Hashimoto K, Kawashima S, Kameoka S, Akiyama Y, Honjaya T, Ejima K, et al. Comparison of image validity between cone beam

Contributors

Feng Ji-ling designed and conducted the experiments, collected part of the CBCT data, analyzed the results, wrote the main manuscript text, and prepared all the figures and tables. Ma Ruo-han, Zhao Jun-ru, Sun Li-li, and Zhao Yan-ping conducted the experiments. Li Gang designed and conducted the experiments and made an essential revision to the manuscript. All the authors reviewed the manuscript.

Funding

This study has received funding by Beijing-Tianjin-Hebei Basic Research Cooperation Project (No.22JCZXC00080), and research foundation of Peking University School and Hospital of Stomatology (PKUSS20220116)

- computed tomography for dental use and Multidetector row Helical computed tomography. *Dentomaxillofac Radiol* 2007; **36**: 465–71. <https://doi.org/10.1259/dmfr/22818643>
12. Ma RH, Yin S, Li G. The detection accuracy of cone beam CT for osseous defects of the Temporomandibular joint: a systematic review and meta-analysis. *Sci Rep* 2016; **6**: 34714. <https://doi.org/10.1038/srep34714>
13. Jie L, Sisi Q, Kaiyuan F. Preliminary study of the stability of Condylar bony changes of the severe joint Osteoarthritis using cone beam computed tomography. *Chin J Orthod* 2017; **244**: 212–16.
14. Koyama J, Nishiyama H, Hayashi T. Follow-up study of Condylar bony changes using Helical computed tomography in patients with Temporomandibular disorder. *Dentomaxillofac Radiol* 2007; **36**: 472–77. <https://doi.org/10.1259/dmfr/28078357>
15. Park SB, Yang YM, Kim YI, Cho BH, Jung YH, Hwang DS. Effect of Bimaxillary surgery on adaptive Condylar head remodeling: metric analysis and image interpretation using cone-beam computed tomography volume Superimposition. *J Oral Maxillofac Surg* 2012; **70**: 1951–59. <https://doi.org/10.1016/j.joms.2011.08.017>
16. Ok S-M, Lee J, Kim Y-I, Lee J-Y, Kim KB, Jeong S-H. Anterior Condylar remodeling observed in stabilization splint therapy for Temporomandibular joint osteoarthritis. *Oral Surg Oral Med Oral Pathol Oral Radiol* 2014; **118**: 363–70. <https://doi.org/10.1016/j.ooolo.2014.05.022>
17. Lee J-H, Lee W-J, Shin J-M, Huh K-H, Yi W-J, Heo M-S, et al. Three-dimensional assessment of Condylar surface changes and remodeling after Orthognathic surgery. *Imaging Sci Dent* 2016; **46**: 25–31. <https://doi.org/10.5624/isd.2016.46.1.25>
18. Jiang YY, Sun L, Wang H, Zhao CY, Zhang WB. Three-dimensional cone beam computed tomography analysis of Temporomandibular joint response to the twin-block functional appliance. *Korean J Orthod* 2020; **50**: 86–97. <https://doi.org/10.4041/kjod.2020.50.2.86>
19. Feng JL, Ma RH, Du H, Zhao YP, Meng JH, Li G. Diagnostic accuracy of fused CBCT images in the evaluation of Temporomandibular joint Condylar bone Resorption. *Clin Oral Invest* 2023; **27**: 1277–88. <https://doi.org/10.1007/s00784-022-04761-x>

20. Schilling J, Gomes LCR, Benavides E, Nguyen T, Paniagua B, Styner M, et al. Regional 3d Superimposition to assess Temporomandibular joint Condylar morphology. *Dentomaxillofac Radiol* 2014; **43**(1): 20130273. <https://doi.org/10.1259/dmfr.20130273>
21. Schiffman E, Ohrbach R, Truelove E, Look J, Anderson G, Goulet J-P, et al. Diagnostic criteria for Temporomandibular disorders (DC/TMD) for clinical and research applications: recommendations of the International RDC/TMD consortium network* and orofacial pain special interest group. *J Oral Facial Pain Headache* 2014; **28**: 6–27. <https://doi.org/10.11607/jop.1151>
22. Ma RH, Li G, Sun Y, Meng JH, Zhao YP, Zhang H. Application of fused image in detecting abnormalities of Temporomandibular joint. *Dentomaxillofac Radiol* 2019; **48**(3): 20180129. <https://doi.org/10.1259/dmfr.20180129>
23. Koo TK, Li MY. A guideline of selecting and reporting Intraclass correlation coefficients for reliability research. *J Chiropr Med* 2016; **15**: 155–63. <https://doi.org/10.1016/j.jcm.2016.02.012>
24. Styner M, Oguz I, Xu S, Brechbühler C, Pantazis D, Levitt JJ, et al. Framework for the statistical shape analysis of brain structures using SPHARM-PDM. *Insight J* 2006; 242–50.
25. Paniagua B, Cevidanes L, Walker D, Zhu H, Guo R, Styner M. Clinical application of SPHARM-PDM to quantify Temporomandibular joint osteoarthritis. *Comput Med Imaging Graph* 2011; **35**: 345–52. <https://doi.org/10.1016/j.compmedimag.2010.11.012>
26. Garcia-Sanz V, Bellot-Arcís C, Hernández V, Serrano-Sánchez P, Guarinos J, Paredes-Gallardo V. Accuracy and reliability of cone-beam computed tomography for linear and volumetric Mandibular Condyle measurements. A human cadaver study. *Sci Rep* 2017; **7**(1): 11993. <https://doi.org/10.1038/s41598-017-12100-4>
27. Wang Y-H, Li G, Ma R-H, Zhao Y-P, Zhang H, Meng J-H, et al. Diagnostic efficacy of CBCT, MRI, and CBCT-MRI fused images in distinguishing Articular disc calcification from loose body of Temporomandibular joint. *Clin Oral Investig* 2021; **25**: 1907–14. <https://doi.org/10.1007/s00784-020-03497-w>
28. Wang Y-H, Ma R-H, Li J-J, Mu C-C, Zhao Y-P, Meng J-H, et al. Diagnostic efficacy of CBCT, MRI and CBCT-MRI fused images in determining anterior disc displacement and bone changes of Temporomandibular joint. *Dentomaxillofac Radiol* 2022; **51**(2): 20210286. <https://doi.org/10.1259/dmfr.20210286>