International Journal of Computer Assisted Radiology and Surgery (2022) 17:751–759
https://doi.org/10.1007/s11548-021-02513-y

ORIGINAL ARTICLE

Validity and reliability of masseter muscles segmentation from the transverse sections of Cone-Beam CT scans compared with MRI scans

Yichen Pan[1,2,3] · Yinghui Wang[4] · Gang Li[4] · Si Chen[1,2,3] · Tianmin Xu[1,2,3]

Received: 14 July 2021 / Accepted: 27 September 2021 / Published online: 8 October 2021

Abstract

Background To evaluate the validity and reliability of cone-beam computed tomography (CBCT) masseter muscle segmentation by comparing with the magnetic resonance imaging (MRI) masseter muscle segmentation of the same patients.

Methods Seventeen volunteers were included in this study. CBCT and MRI scans of the volunteers were taken, respectively, within one month. The masseter muscles in the CBCT scans were segmented by a generative adversarial network (GAN)-based framework combined with manual check. The masseter muscles in the MRI scans were segmented manually. The segmentations were repeated by the first examiner and a second examiner. For cross-sectional area (CSA), paired t-test, intraclass correlation coefficient (ICC) and standard error of measurement (SEM) were calculated to evaluate the validity and reliability of the segmentations. The validity and reliability were also calculated by Dice similarity coefficient (DSC) and average Hausdorff distance (aHD) between different segmentations.

Seventeen volunteers were included in this study. CBCT and MRI scans of the volunteers were taken, respectively, within one month. The masseter muscles in the CBCT scans were segmented by a generative adversarial network (GAN)-based framework combined with manual check. The masseter muscles in the MRI scans were segmented manually. The segmentations were repeated by the first examiner and a second examiner. For cross-sectional area (CSA), paired t-test, intraclass correlation coefficient (ICC) and standard error of measurement (SEM) were calculated to evaluate the validity and reliability of the segmentations. The validity and reliability were also calculated by Dice similarity coefficient (DSC) and average Hausdorff distance (aHD) between different segmentations.

Results Paired t-test showed that there was no significant difference in CSA between CBCT and MRI masseter segmentations. The ICCs were all larger than 0.95 and the SEM was less than 4.85 mm^2 for CSA. The DSC was all larger than 0.95 showing over 95% of similarity between CBCT and MRI masseter segmentations. The aHD was all smaller than 0.09 mm showing great consistency of the contour of CBCT and MRI segmentations.

Conclusion Masseter muscle segmentation from CBCT scans was not significantly different from the segmentation from MRI scans. CBCT muscle segmentation showed great validity compared with MRI scans, and great reliability in retests.

Keywords Cone-beam computed tomography · Magnetic resonance imaging · Masseter muscle · Computer-assisted image processing · Reliability and validity

✉ Si Chen
elisa02@163.com

✉ Tianmin Xu
tmxuortho@163.com

[1] Department of Orthodontics, Peking University School and Hospital of Stomatology, 22 Zhongguancun Avenue South, Haidian District, 100081 Beijing, People's Republic of China

[2] National Engineering Laboratory for Digital and Material Technology of Stomatology, 22 Zhongguancun Avenue South, Haidian District, 100081 Beijing, People's Republic of China

[3] Beijing Key Laboratory of Digital Stomatology, 22 Zhongguancun Avenue South, Haidian District, Beijing 100081, People's Republic of China

[4] Department of Oral and Maxillofacial Radiology, Peking University School and Hospital of Stomatology, 22 Zhongguancun Avenue South, Haidian District, Beijing 100081, People's Republic of China

Abbreviations

CBCT	Cone-beam computed tomography
MRI	Magnetic resonance imaging
ICC	Intraclass correlation coefficient
SEM	Standard error of measurement
CSA	Cross-sectional area
DSC	Dice similarity coefficient
HD	Hausdorff distance
aHD	Average Hausdorff distance

Background

Orthodontic tooth movement was not only affected by mechanical force exerted by wire-bracket system, but also by physiological forces generated by the orofacial system [1]. The occlusal force generated by the masticatory muscles is an important component of the physiological forces. The masseter muscles are large and superficial [2] which greatly contribute to the occlusal force and serves as an important role in mastication activity which closely correlates with the long-term morphology of the facial type (dolichofacial, mesofacial and brachyfacial) [3], temporomandibular joints(TMJ) [4, 5], stability of the dental arch [6], etc. Furthermore, the masseter muscles also support the facial soft tissue around the mandibular gonion area which is crucial to frontal facial profile. While esthetics have gradually become a major concern of the orthodontic patients and research shows that posttreatment decrease in the facial width [7, 8] may be related to atrophy of the masseter muscle [9], it is important to study the masseter muscle in orthodontic area.

Different medical imaging methods are being used to study the masseter muscle. Ultrasound was used to measure the thickness of the masseter muscle [10, 11]. CT and MRI can both clearly display soft tissue and have been widely used in studying muscle's cross-sectional areas or volumes [12–14]. High consistency has been found between CT and MRI in the cross sections of the masseter muscle and the medial pterygoid muscle [13]. Furthermore, the image of muscles in CT and MRI scans have also been confirmed to be consistent with dissections [13, 15]. However, CBCT scanning is the only widely used 3-dimentional medical imaging technique in orthodontic area, while MRI and CT are not regular tests. The metal restorations or brackets will produce significant artifacts in MRI scans, not to mention the cost and time needed for a single scan. Large radiation dose of CT limits its use for non-surgical patients in dental practice.

Cone-beam computed tomography (CBCT) has been widely used in orthodontic area for three-dimensional measurements. CBCT can achieve less slice thickness with much lower radiation dose [16]. Highly improved resolution of the advanced CBCT technology makes it possible for the masseter muscle to be distinguished from CBCT scans [17] and the transverse sections of CBCT scans can be used to measure the cross-sectional area of the masseter muscle [18]. MRI is the widely recognized and examined imaging method for muscle tissues. However, to our knowledge, there's no quantitative study comparing the measurement of masseter muscle from CBCT scans with MRI scans. The gray scale of soft tissues in CBCT was nonlinear and within a limited range, making it difficult for auto-segmentation based on the threshold value. Manually segmentation of the masseter muscles from the CBCT scans is tedious and time-consuming because hundreds of layers need to be processed. Fortunately, due to the development of artificial intelligence technology, automatic medical image procession has been greatly improved. In combination of machine learning, automatic masseter muscles segmentation has been realized on both CT [19, 20] and CBCT scans [21]. Therefore, the purpose of this study was to evaluate the validity and reliability of the masseter muscle segmentation from the transverse sections of CBCT scans compared with MRI scans.

Methods

Acquisition and registration of CBCT and MRI scans

The volunteers were selected from the sequential patients seeking for treatment in the department of orthodontics, Peking University School and Hospital of Stomatology. The imaging data of the patients will be used for a further study on temporomandibular joint (TMJ). The inclusion and exclusion criteria were as follows: (1) 18–35 years old, (2) with complete permanent teeth (excluding the third molars) and without posterior buccal and lingual crossbites, (3) without craniofacial congenital diseases or soft tissue diseases, (4) without craniofacial trauma, tumors or history of surgery, (5) without history of orthodontic treatment, (6) without intraoral metal restorations or prostheses, (7) without any contraindications for MRI scanning.

A total of 20 volunteers were included in this study. But 3 were removed due to the deformed images that resulted from the movement in the scanning procedure of MRI, and at last, 17 patients were included in this study. Each volunteer accepted CBCT and MRI scanning at an interval of less than one month and ensured that the body weight change was less than 1 kg. Each volunteer signed a written informed consent. And the protocol was reviewed and approved by the Institutional Review Board of Peking University School and Hospital of Stomatology (PKUSSIRB-201944062).

CBCT scans were taken by Newtom VGi (Quantitative Radiology, Verona, Italy) with the following settings: field of view (FOV) = 15 × 15 cm, 110 kV; scan time = 27 s; voxel

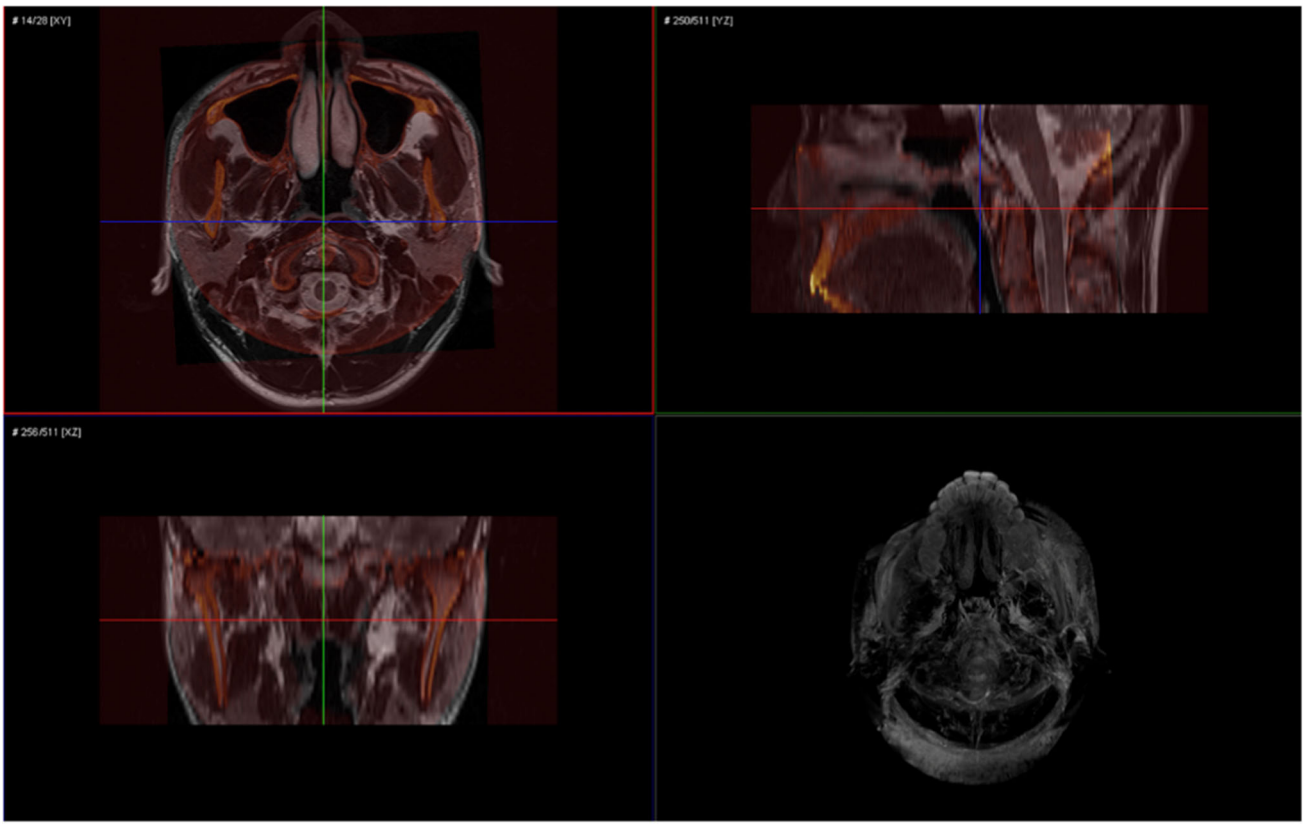

Fig. 1 Registration result of one of the volunteers: top right shows the selection of upper layer for segmentation defined as the axial section passing through the tip of nose; the middle and lower layer were defined as the axial scan 9 mm and 18 mm lower than the upper layer, respectively

size = 0.2 mm. T2-weighed MRI scans were performed on 3.0 T MAGNETOM Trio Tim (Siemens Healthcare, Erlangen, Germany) scanner with the following settings: repetition time (TR) = 3700, echo time (TE) = 90, flip angle = 120, field of view (FOV) = 512×512; slice thickness = 3 mm; spacing between slices = 3.6 mm, acquisition time $\approx$ 20 min.

CBCT and MRI data in Digital Imaging and Communications in Medicine (DICOM) format of each volunteer were imported to Amira visual software (version 5.4.3, Visage Imaging, Melbourne, Australia) and registered through its inherent function module (Fig. 1) [22]. In our study, the MRI scans were fixed and CBCT scans were registered onto the MRI scans.

Masseter muscle segmentation in transverse sections

Since the orientation of original CBCT and MRI scans was different, CBCT scans were resliced by the Resample function module in the Amira software using MRI transverse scans as a reference.

In order to represent the upper, middle, lower parts of the masseter muscle, three layers in each patient were selected for segmentation and comparison. Layers for masseter muscle segmentation were selected from the sagittal view (Fig. 1, top right window). The upper layer was defined as the axial

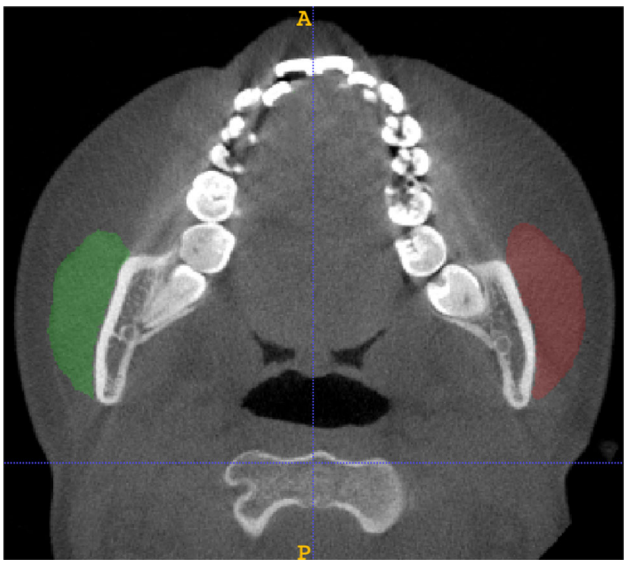

Fig. 2 One layer of CBCT scans and masseter muscle segmentation (green and red) from the ITK-SNAP software

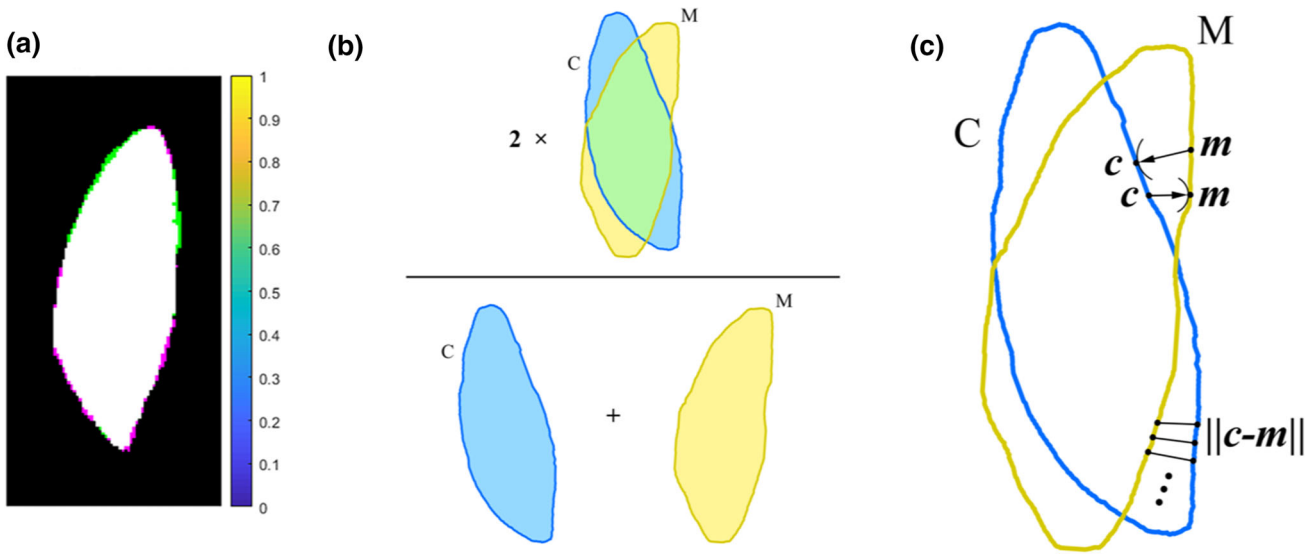

Fig. 3 The deviation between correspondent CBCT and MRI segmentations **a** General deviation between two segmentations showed by color bar;). **b** DSC calculation: C (blue) for CBCT segmentation and M (yellow) for MRI segmentation. C ∩ M is the overlapped area (green) between them, and C + M is the total area (blue and yellow) of the two segmentations. **c** aHD calculation: arrows connect the paired points ($\boldsymbol{c} \rightarrow \boldsymbol{m}$ or $\boldsymbol{m} \rightarrow \boldsymbol{c}$) between the two segmentations, and $\|\boldsymbol{c}\text{-}\boldsymbol{m}\|$ means the distance between $\boldsymbol{c}$ and $\boldsymbol{m}$

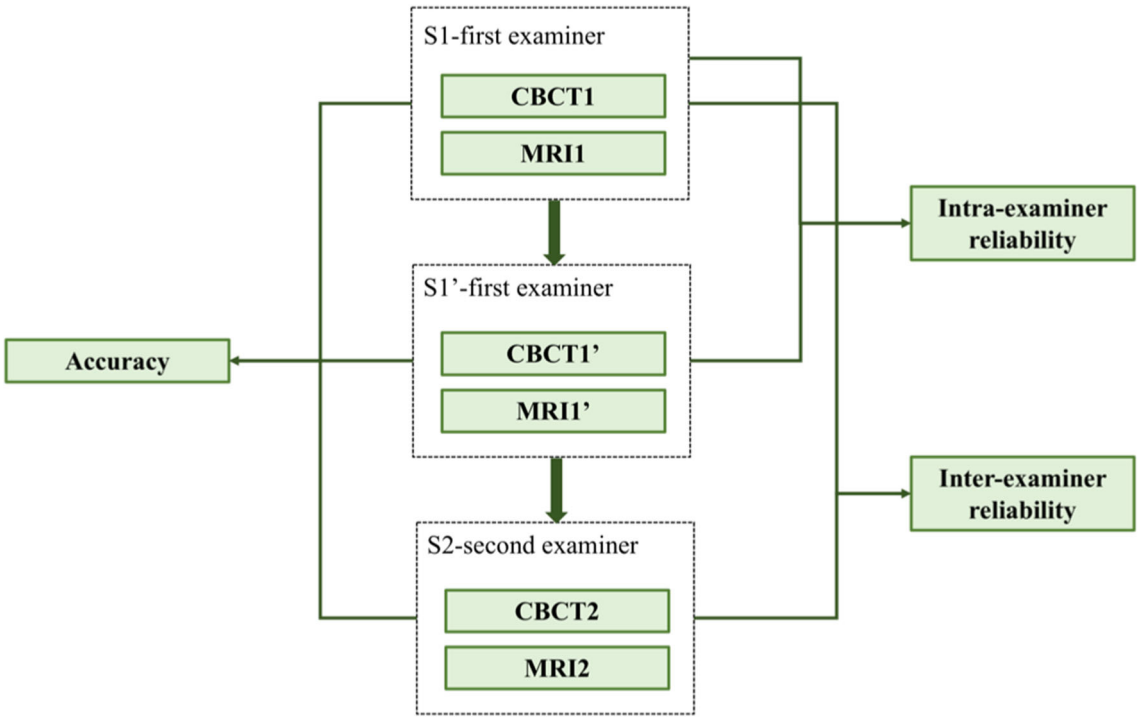

Fig. 4 Flowchart of the study showing the labels of different segmentations and comparisons

section passing through the tip of nose. The middle and lower layer were defined as the axial scan 9 mm and 18 mm lower than the upper layer respectively. Because the MRI scans mainly focused on TMJ, not all the layers of masseter muscle were shown in all the patients, the number of layers will be clarified in the following tables. Left and right sides of each patient were regarded as two samples, so in total, 34 masseter muscles in the 17 patients were included.

A self-developed generative adversarial network (GAN)-based framework [21] was used for noise reduction and automatic segmentation of masseter muscles from the transverse sections of CBCT scans. The framework was developed by the Department of Machine Intelligence, Key Laboratory of Machine Perception (MOE), Peking University. A layer by layer manual check was performed using ITK-SNAP 3.6.0 (http://www.itksnap.org) based on the automatic segmentation result (Fig. 2). For MRI scans, the segmentation was manually performed using ITK-SNAP 3.6.0 directly.

Comparisons between CBCT and MRI masseter muscles

The segmentation result was exported for the analysis of validity and reliability of the masseter muscle segmentation.

Cross-sectional area (CSA) of each layer of both CBCT and MRI scans was calculated for the evaluation of validity and reliability of masseter muscle segmentation.

Since bias cannot be ruled out even the cross-sectional areas were consistent, we introduced Dice Similarity Coefficient (DSC)[23] and average Hausdorff distance (aHD) [24] in order to detect the deviation between the morphology and position of the corresponding CBCT and MRI segmentations. The segmentations were exported from ITK-SNAP 3.6.0 software (Fig. 3a) and DSC and aHD were calculated by self-compiled scripts.

The DSC indicates the overlap between the CBCT and MRI muscle segmentation (Fig. 3b) and was defined as follows:

$$\mathrm{DSC} = \frac{2\mathrm{C} \cap \mathrm{M}}{(\mathrm{C} + \mathrm{M})} \tag{1}$$

(C was the CBCT segmentation and M as the MRI segmentation, $\mathrm{C} \cap \mathrm{M}$ was the overlap between them, and the $\mathrm{C} + \mathrm{M}$ was the total area of the two segmentations.)

The aHD was modified from the Hausdorff Distance (HD) [25] and calculates the average distance between the CBCT segmentation and the MRI segmentation. A point on the contour of the CBCT segmentation and its nearest point on the contour of the MRI segmentation or vice versa form a point pair ($\boldsymbol{c} \rightarrow \boldsymbol{m}$ or $\boldsymbol{m} \rightarrow \boldsymbol{c}$ in Fig. 3c). However, the Hausdorff Distance (HD) measures the largest distance between the point pairs ($\max(\|\boldsymbol{c}\text{-}\boldsymbol{m}\|)$) and therefore is very sensible to outliers, so we used the average Hausdorff Distance (aHD) [24] defined as follows to measure the average deviation between two segmentations (Fig. 3c):

$$\mathrm{ahd}(\mathrm{C}, \mathrm{M}) = \frac{1}{|\mathrm{C}|} \sum_{\boldsymbol{c} \in \mathrm{C}} \min_{\boldsymbol{m} \in \mathrm{M}} (\|c - m\|) \tag{2}$$

$$\mathrm{aHD} = \frac{1}{2}(\mathrm{ahd}(\mathrm{C}, \mathrm{M}) + \mathrm{ahd}(\mathrm{M}, \mathrm{C})) \tag{3}$$

(C is the contour of CBCT segmentation and M the MRI. $\boldsymbol{c}$ is a point on C, and $\boldsymbol{m}$ on M. |C| is the total number of $\boldsymbol{c}$. $\|\boldsymbol{c}\text{-}\boldsymbol{m}\|$ means the distance between c and m. Equation (2) tells that, for every $\boldsymbol{c}$ of C, we find the nearest point $\boldsymbol{m}$ of M, then take the average over all of $\boldsymbol{c}$'s. The equations refer to Alba et al.[24].)

Statistical analysis

The segmentation of the 34 masseter muscles (S1) was repeated by the same examiner at least two weeks later (S1') and the second examiner (S2) (Fig. 4). Validity was evaluated by comparisons between CBCT and MRI segmentations (CBCT1-MRI1, CBCT1'-MRI1', CBCT2-MRI2). Intra-examiner reliability of the muscle segmentation was calculated by the difference between S1 and S1' (CBCT1-CBCT1', MRI1-MRI1'), and inter-examiner reliability of the muscle segmentation was calculated by the difference between S1 and S2 (CBCT1-CBCT2, MRI1-MRI2) (Fig. 4).

For CSA, paired t-test, intraclass correlation coefficient (ICC) and Standard error of measurement (SEM) [26] were calculated to evaluate the validity and reliability of the segmentations. SEM was calculated as follows[26]:

$$SEM = S_x \sqrt{1 - ICC} \tag{4}$$

(S_x is the pooled standard deviation.)

DSC and aHD were directly calculated between different segmentations to evaluate the validity and reliability.

The statistical analysis of CSA was performed using SPSS 27.0 (IBM, Armonk, N.Y.), with a significance level of 0.05.

Results

Descriptive statistics

Finally, 34 masseter muscles of 17 patients (4 males and 13 females) were analyzed in this study. The average age of the sample was 24.80 ± 3.56 years old and their mean body mass index (BMI) was 21.23 kg/m^2.

Cross-sectional area (CSA)

The validity of the CSA was compared by the difference in CSA between CBCT and MRI scans in each segmentations (CBCT1-MRI1, CBCT1'-MRI1', CBCT2-MRI2). The difference in CSAs of the three layers between the three CBCT and MRI segmentations was listed in the first three rows of Table 1. All the differences in CSAs were not statistically significant. The reliability of CSA was represented by the difference in CBCT or MRI segmentations between S1 and

Table 1 Paired t-test of the difference of CSA (mm^2) with zero between different scans of masseter muscle

		Upper		Middle		Lower	
		$N = 16 \times 2$*		$N = 13 \times 2$		$N = 11 \times 2$	
		Mean ± SD	p value	Mean ± SD	p value	Mean ± SD	p value
Validity	CBCT1-MRI1	− 1.70 ± 19.22	0.80	− 1.52 ± 28.54	0.94	− 8.57 ± 22.49	0.09
	CBCT1'- MRI1'	− 5.91 ± 24.81	0.08	1.93 ± 31.35	0.95	− 8.27 ± 19.44	0.06
	CBCT2-MRI2	− 2.35 ± 24.48	0.66	− 4.14 ± 24.71	0.43	− 8.31 ± 21.49	0.08
Intra-examiner reliability	CBCT1-CBCT1'	0.90 ± 16.34	0.33	1.83 ± 18.26	0.30	2.47 ± 18.70	0.54
	MRI1-MRI1'	− 3.86 ± 15.92	0.29	4.54 ± 21.50	0.17	− 8.04 ± 24.00	0.13
Inter-examiner reliability	CBCT1-CBCT2	− 3.30 ± 21.57	0.29	5.28 ± 13.85	0.12	2.78 ± 22.86	0.58
	MRI1-MRI2	− 4.51 ± 19.54	0.25	1.92 ± 17.19	0.53	− 7.78 ± 29.63	0.23

*$N = 16 \times 2$ means that 16 patients with upper layer times 2 sides of each patient (same in middle and lower layers and in the following tables)

Table 2 ICC and SEM of the CSA (mm^2) measurements

		ICC			SEM (mm^2)		
		Upper $N = 16 \times 2$	Middle $N = 13 \times 2$	Lower $N = 11 \times 2$	Upper $N = 16 \times 2$	Middle $N = 13 \times 2$	Lower $N = 11 \times 2$
Validity	CBCT1-MRI1	0.97	0.95	0.98	3.18	4.85	1.90
	CBCT1'-MRI1'	0.98	0.95	0.97	2.09	4.72	2.84
	CBCT2-MRI2	0.98	0.97	0.98	2.19	2.80	2.05
Intra-examiner reliability	CBCT1-CBCT1'	0.98	0.98	0.99	2.12	2.02	0.96
	MRI1-MRI1'	0.97	0.99	0.98	3.15	0.90	1.87
Inter-examiner reliability	CBCT1-CBCT2	0.97	0.98	0.99	3.34	1.95	1.00
	MRI1-MRI2	0.96	0.98	0.98	4.16	1.85	1.95

Table 3 Comparison of dice similarity coefficient (DSC) of masseter muscles segmentation from CBCT and MRI scans and remeasurements

		Upper $N = 16 \times 2$	Middle $N = 13 \times 2$	Lower $N = 11 \times 2$
Validity	CBCT1-MRI1	0.95 ± 0.02	0.96 ± 0.02	0.94 ± 0.01
	CBCT1'- MRI1'	0.96 ± 0.02	0.95 ± 0.02	0.95 ± 0.02
	CBCT2-MRI2	0.95 ± 0.02	0.96 ± 0.01	0.95 ± 0.03
Intra-examiner reliability	CBCT1-CBCT1'	0.97 ± 0.02	0.97 ± 0.01	0.97 ± 0.02
	MRI1-MRI1'	0.97 ± 0.02	0.97 ± 0.02	0.96 ± 0.02
Inter-examiner reliability	CBCT1-CBCT2	0.97 ± 0.02	0.97 ± 0.01	0.96 ± 0.01
	MRI1-MRI2	0.97 ± 0.02	0.97 ± 0.01	0.96 ± 0.02

Table 4 Comparison of average Hausdorff distance (aHD) (mm) of masseter muscles segmentation from CBCT and MRI scans and remeasurements

		Upper $N = 16 \times 2$	Middle $N = 13 \times 2$	Lower $N = 11 \times 2$
Validity	CBCT1-MRI1	0.09 ± 0.21	0.05 ± 0.08	0.06 ± 0.07
	CBCT1'- MRI1'	0.04 ± 0.05	0.03 ± 0.02	0.03 ± 0.02
	CBCT2-MRI2	0.08 ± 0.14	0.06 ± 0.12	0.08 ± 0.12
Intra-examiner reliability	CBCT1-CBCT1'	0.07 ± 0.20	0.04 ± 0.08	0.04 ± 0.07
	MRI1-MRI1'	0.07 ± 0.14	0.05 ± 0.11	0.06 ± 0.11
Inter-examiner reliability	CBCT1-CBCT2	0.04 ± 0.09	0.02 ± 0.01	0.03 ± 0.01
	MRI1-MRI2	0.04 ± 0.11	0.03 ± 0.09	0.05 ± 0.09

S1', S1 and S2 (CBCT1-CBCT1', MRI1-MRI1', CBCT1-CBCT2, MRI1-MRI2) (Table 1, the last four rows).

Intraclass correlation coefficient (ICC) and SEMs between the three CBCT and MRI segmentations were shown in Table 2. In general, the ICCs were higher and SEMs were lower in the second segmentations showing greater validity in the second segmentation than in the first. ICCs were all higher than 0.95. And the SEMs ranged from 1.00 to 4.85 mm^2 in all layers which is relatively small clinically.

Dice similarity coefficient (DSC) and average Hausdorff distance (aHD)

The segmentations of CBCT and MRI scans were mutually compared and DSC and aHD were calculated (Tables 3,4). Validity was represented by the DSC and aHD between CBCT and MRI scans in each segmentation (CBCT1-MRI1, CBCT1'-MRI1', CBCT2-MRI2, first three rows of Tables 3, 4). Reliability was represented by the DSC and aHD between different times of CBCT and MRI segmentations (CBCT1-CBCT1', MRI1-MRI1', CBCT1-CBCT2, MRI1-MRI2, last four rows of Tables 3, 4). The DSCs were all larger than 0.95 which is quite well in segmentation and pretty low aHDs (less than 0.1 mm) were found in all the comparisons.

Discussion

In this study, we found that masseter muscle could be properly displayed in CBCT scans with similar image quality as that in MRI scans. The morphology and contour of masseter muscle segmentations from CBCT scans are not significantly different from the MRI scans.

The morphology and function of masseter muscles are closely related to the dentofacial system. In terms of therapeutic decision, patients with short face are usually accompanied by strong masseter strength, and the alveolar bones could be tougher than those with long face and weak masseter muscle strength [27]. For example, in clinical decision, nonextraction treatment was usually preferred in patients with short face (smaller mandibular angle) since the tooth movement were harder and slower in those with tough alveolar bones [28]. Segmentation and reconstruction of masseter muscles, and furthermore, temporalis and medial and lateral pterygoid muscles, would be helpful in retrospective studies focusing on the relationship between masticatory activities [29], the morphology of mandibles and dental arch. The segmentation of masticatory muscles may also be of great help in the construction of a finite element model of the whole craniofacial system. Usually, this kind of studies were based on CT scans [29, 30]. If the segmentation of masticatory muscles could be realized in CBCT scans, radiation dose exposed to patients could be largely decreased.

MRI is one of the most generally accepted imaging methods to study the morphology of masseter muscles with the advantages of no radiation doses and clear soft tissue display [31]. Studies have been done on the soft tissue clearness in CT and MRI, which have found that CT have similar clearance in studying muscle tissues to MRI [13] and both are in great concordance with anatomic dissections [13, 15, 32]. However, due to previously mentioned limitations, neither of the two imaging methods were widely used in orthodontic area. CBCT, instead, is becoming a common imaging method in orthodontic practice [33]. Therefore, in this study, we would like to explore that whether the masseter muscle image in CBCT can also be used for soft tissue analysis.

Three layers relative to the masseter muscle were chosen for segmentation in this study, because different structures may influence the segmentation of masseter muscles in CBCT scans. The parotid gland was located posterior, and the buccal fat pad was located lateral and anterior to the masseter muscle [2], and they were all of similar gray scale in CBCT scans which may be hard to distinguish with the muscle tissue [34].

For CSA, the validity was represented by the consistency between CBCT and MRI scans of the same patients, with MRI being used as the "golden standard" and the reliability was represented by the consistency between multiple measurements. Usually, the consistency was compared using two categories of indicators, correlation analysis and error analysis. Correlation analysis includes Pearson's correlation coefficient [35, 36], intraclass correlation coefficient (ICC) [12] and regression analysis [15]. Error analysis includes standard error [12] or method error. While correlation analysis reflects the variance ranging from 0 to 1, the standard error has the same units with the measurements and is not influenced by the variability of the patients [12, 26]. Therefore, in this study, we choose both ICC and SEM. The ICC was interpreted as follows: below 0.699 was poor, 0.700–0.799 was fair, 0.800–0.899 was good, and 0.900 to 1.000 was excellent [12]. In this study, the ICCs were all larger than 0.95 which was excellent consistency. The SEM for CSA was 1.00–4.85 mm^2 which was clinically acceptable.

In terms of the selection of comparison variables, DSC reflects the internal characteristics of the morphology [23] and HD reflects the marginal characteristics of the contour [25]. However, it is well-known that HD could be easily influenced by noise and extreme value of the segmentation margins. Therefore, to reduce the influence of outliers, we used a modified parameter—the average Hausdorff distance (aHD) in this study to measure the average deviation between two segmentations [24]. Noise could be erased in this study because all the CBCT segmentations was manually checked and corrected and all the MRI segmentations were depicted manually. In this study, high DSC (over 0.95) and low aHD

(under 0.1 mm) shows great morphological similarity and reliability of CBCT and MRI segmentations.

In the actual clinical practice, it is impossible to obtain the real cross sections of the patients' masseter muscles, so MRI was one of the best approaches to study the morphology of masseter muscles. If there was no significant difference in the morphology of masseter muscles between CBCT and MRI, it may be the basis for us to replace MRI by CBCT in studying masseter muscles. Our study showed that CBCT scans taken with proper settings can replace MRI in studying the morphology of masseter muscles with both great validity and reliability.

There are also several limitations in this study. Motions during the MRI scanning may influence the trueness of the morphology of masseter muscles. Different scanning position may lead to some extent of soft tissue deformation between MRI (supine position) and CBCT (upright position). In addition, although the patients were asked to keep in light occlusal position during the imaging process, it could also be possible that the patients changed to rest position instead during the long duration of MRI scanning.

The advantage of CBCT is shorter scan time, smaller slice thickness and no slice gap [33]. Because of short scan time, it is easier for the patients to keep occlusal position without movement, which decrease the possibility of deformation due to movement of the patient. Also, CBCT scans used in orthodontic area can have voxel resolutions ranging from 0.125 to 0.4 mm [33] which makes it more capable of three-dimensional reconstruction. With similar segmentation outcomes, CBCT could serve as a more convenient and operationally controllable approach for muscle segmentation tasks.

Conclusions

Masseter muscle segmentation from CBCT scans taken with proper settings was not significantly different from the segmentation from MRI scans. CBCT muscle segmentation showed great validity compared with MRI scans, and great reliability in retests. Therefore, masseter muscle images in CBCT scans can be a source of three-dimensional information for orthodontic clinic and research.

Acknowledgements We thank Yungeng Zhang and Yuru Pei from Key Laboratory of Machine Perception (MOE), Department of Machine Intelligence, Peking University for the technical support in CBCT masseter muscle autosegmentation for this study.

Authors' contributions SC, GL and TX designed the study together. YP, YW carried out the data collection, the measurement and remeasurements, analyzed the data and prepared the tables and figures. YP, YW and SC discussed the results and YP drafted the manuscript. SC, GL and TX critically reviewed the manuscript. All authors have read and approved the final version of the manuscript.

Funding This work was supported by the National Natural Science Foundation of China (81671034), National Natural Science Foundation of China (81200806), Beijing Natural Science Foundation (7192227) and Peking University Medicine Seed Fund for Interdisciplinary Research (BMU 2018MI013). The authors declare no conflicts of interest in this study.

Declarations

Conflict of interests The authors declare that they have no competing interests.

Ethics approval The study was reviewed and approved by the Institutional Review Board of Peking University School and Hospital of Stomatology (PKUSSIRB-201944062).

Consent to participate Written informed consent was obtained from each patient before participation in the study.

Availability of data and materials The full datasets used and analyzed during the current study are available on reasonable request from the corresponding authors at tmxuortho@163.com and elisa02@163.com.

References

1. Teng F, Du FY, Chen HZ, Jiang RP, Xu TM (2019) Three-dimensional analysis of the physiologic drift of adjacent teeth following maxillary first premolar extractions. Sci Rep 9(1):14549
2. Gaudy JF, Zouaoui A, Bravetti P, Charrier JL, Guettaf A (2000) Functional organization of the human masseter muscle. Surg Radiol Anat 22(3–4):181–190
3. Becht MP, Mah J, Martin C, Razmus T, Gunel E, Ngan P (2014) Evaluation of masseter muscle morphology in different types of malocclusions using cone beam computed tomography. Int Orthod 12(1):32–48
4. He S, Wang S, Song F, Wu S, Chen J, Chen S (2021) Effect of the use of stabilization splint on masticatory muscle activities in TMD patients with centric relation-maximum intercuspation discrepancy and absence of anterior/lateral guidance. Cranio 39(5):424–432
5. Nickel JC, Weber AL, Covington Riddle P, Liu Y, Liu H, Iwasaki LR (2017) Mechanobehaviour in dolichofacial and brachyfacial adolescents. Orthod Craniofac Res 20(Suppl 1):139–144
6. Farronato G, Giannini L, Galbiati G, Stabilini SA, Sarcina M, Maspero C (2015) Functional evaluation in orthodontic surgical treatment: long-term stability and predictability. Prog Orthod 16:30
7. Pan Y, Chen S, Shen L, Pei Y, Zhang Y, Xu T (2020) Thickness change of masseter muscles and the surrounding soft tissues in female patients during orthodontic treatment: a retrospective study. BMC Oral Health 20(1):181
8. Dai F, Yu J, Chen G, Xu T, Jiang R (2018) Changes in buccal facial depth of female patients after extraction and nonextraction orthodontic treatments: A preliminary study. Korean J Orthod 48(3):172–181
9. Kiliaridis S, Mills CM, Antonarakis GS (2010) Masseter muscle thickness as a predictive variable in treatment outcome of the twin-block appliance and masseteric thickness changes during treatment. Orthod Craniofac Res 13(4):203–213
10. Busato A, Balconi G, Vismara V, Bertele L, Garo G, Deg D (2016) Management and control of isotonic contraction generated stress: evaluation of masseter muscle deformation pattern by means of ecography. Oral Implantol (Rome) 9(Suppl 1/2016 to N 4/2016):45–53

11. Lione R, Franchi L, Noviello A, Bollero P, Fanucci E, Cozza P (2013) Three-dimensional evaluation of masseter muscle in different vertical facial patterns: a cross-sectional study in growing children. Ultrason Imaging 35(4):307–317
12. Hu ZJ, He J, Zhao FD, Fang XQ, Zhou LN, Fan SW (2011) An assessment of the intra- and inter-reliability of the lumbar paraspinal muscle parameters using CT scan and magnetic resonance imaging. Spine (Phila Pa 1976) 36(13):E868–E874
13. van Spronsen PH, Weijs WA, Valk J, Prahl-Andersen B, van Ginkel FC (1989) Comparison of jaw-muscle bite-force cross-sections obtained by means of magnetic resonance imaging and high-resolution CT scanning. J Dent Res 68(12):1765–1770
14. Lee YH, Lee KM, Auh QS (2021) MRI-based assessment of masticatory muscle changes in TMD patients after whiplash injury. J Clin Med 10(7)
15. Mitsiopoulos N, Baumgartner RN, Heymsfield SB, Lyons W, Gallagher D, Ross R (1998) Cadaver validation of skeletal muscle measurement by magnetic resonance imaging and computerized tomography. J Appl Physiol (1985) 85(1):115–122
16. Engstrom CM, Loeb GE, Reid JG, Forrest WJ, Avruch L (1991) Morphometry of the human thigh muscles. A comparison between anatomical sections and computer tomographic and magnetic resonance images. J Anat 176:139–156
17. Demehri S, Muhit A, Zbijewski W, Stayman JW, Yorkston J, Packard N, Senn R, Yang D, Foos D, Thawait GK, Fayad LM, Chhabra A, Carrino JA, Siewerdsen JH (2015) Assessment of image quality in soft tissue and bone visualization tasks for a dedicated extremity cone-beam CT system. Eur Radiol 25(6):1742–1751
18. Lee HJ, Kim SJ, Lee KJ, Yu HS, Baik HS (2017) Repeated injections of botulinum toxin into the masseter muscle induce bony changes in human adults: A longitudinal study. Korean J Orthod 47(4):222–228
19. Chen W, Li Y, Dyer BA, Feng X, Rao S, Benedict SH, Chen Q, Rong Y (2020) Deep learning vs. atlas-based models for fast auto-segmentation of the masticatory muscles on head and neck CT images. Radiat Oncol 15(1):176
20. Zhang X, Chen H, Chen W, Dyer BA, Chen Q, Benedict SH, Rao S, Rong Y (2020) Technical note: atlas-based auto-segmentation of masticatory muscles for head and neck cancer radiotherapy. J Appl Clin Med Phys 21(10):233–240
21. Zhang Y, Pei Y, Qin H, Guo Y, Ma G, Xu T, Zha H (2019) Masseter Muscle Segmentation from Cone-Beam CT Images using Generative Adversarial Network. In IEEE 16th international syposium on biomedical imaging; Venice, Italy: IEEE
22. Ma RH, Li G, Sun Y, Meng JH, Zhao YP, Zhang H (2019) Application of fused image in detecting abnormalities of temporomandibular joint. Dentomaxillofac Radiol 48(3):20180129
23. Dice LR (1945) Measures of the amount of ecologic association between species. Ecology 26(3):297–302
24. Alba JL, Pujol AP, Villanueva JJ (2001) ST-SOM: A shape+texture welf organizing map. The IX Spanish Symposium on Pattern Recognition and Image Analysis; 2001 16-18 May; Benicasim (Castellón), Spain: Universitat Jaume I
25. Huttenlocher DP, Klanderman GA, Rucklidge WJ (1993) Comparing images using the Hausdorff distance. IEEE Trans Pattern Anal Mach Intell 15(9):850–863
26. Ranson CA, Burnett AF, Kerslake R, Batt ME, O'Sullivan PBJESJ (2006) An investigation into the use of MR imaging to determine the functional cross sectional area of lumbar paraspinal muscles. Eur Spine J 15(6):764–773
27. Chaturvedi S, Alfarsi MA (2019) 3-D mapping of cortical bone thickness in subjects with different face form and arch form: A CBCT analysis. Niger J Clin Pract 22(5):616–625
28. Kau CH, Cruz Wilma DA (2020) 3D analysis of tooth movement using 3D technology. Curr Osteoporos Rep. 2020 Oct 10
29. Dai F, Wang L, Chen G, Chen S, Xu T (2016) Three-dimensional modeling of an individualized functional masticatory system and bite force analysis with an orthodontic bite plate. Int J Comput Assist Radiol Surg 11(2):217–229
30. Sana S, Kondody RT, Talapaneni AK, Fatima A, Bangi SL (2021) Occlusal stress distribution in the human skull with permanent maxillary first molar extraction: A 3-dimensional finite element study. Am J Orthod Dentofacial Orthop 14;S0889–5406(21)00404–2
31. Gaudino C, Cosgarea R, Heiland S, Csernus R, Beomonte Zobel B, Pham M, Kim TS, Bendszus M, Rohde S (2011) MR-Imaging of teeth and periodontal apparatus: an experimental study comparing high-resolution MRI with MDCT and CBCT. Eur Radiol 21(12):2575–2583
32. Weijs WA, Hillen B (1984) Relationships between masticatory muscle cross-section and skull shape. J Dent Res 63(9):1154–1157
33. Macleod I, Heath N (2008) Cone-beam computed tomography (CBCT) in dental practice. Dent Update 35(9):590–598
34. Spin-Neto R, Gotfredsen E, Wenzel A (2013) Impact of voxel size variation on CBCT-based diagnostic outcome in dentistry: a systematic review. J Digit Imaging 26(4):813–820
35. Nie M, Liu C, Pan YC, Jiang CX, Li BR, Yu XJ, Wu XY, Zheng SN (2018) Development and evaluation of oral Cancer quality-of-life questionnaire (QOL-OC). BMC Cancer 18(1):523
36. Munawar K, Aqeel M, Rehna T, Shuja KH, Bakrin FS, Choudhry FR (2021) Validity and Reliability of the Urdu Version of the McLean Screening Instrument for Borderline Personality Disorder. Front Psychol 12:533526