Computer-assisted assessment of segmental bimaxillary surgery using voxel- and surface-based registration

A comparative study

Holte, Michael Boelstoft; Diaconu, Alexandru; Pinholt, Else Marie

Published in:
Advances in Oral and Maxillofacial Surgery

DOI:
10.1016/j.adoms.2023.100470

Publication date:
2024

Document version:
Final published version

Citation for pulished version (APA):
Holte, M. B., Diaconu, A., & Pinholt, E. M. (2024). Computer-assisted assessment of segmental bimaxillary surgery using voxel- and surface-based registration: A comparative study. *Advances in Oral and Maxillofacial Surgery*, *13*, Article 100470. https://doi.org/10.1016/j.adoms.2023.100470

Advances in Oral and Maxillofacial Surgery 13 (2024) 100470

Contents lists available at ScienceDirect

Advances in Oral and Maxillofacial Surgery

journal homepage: www.sciencedirect.com/journal/advances-in-oral-and-maxillofacial-surgery

Full Length Article

Computer-assisted assessment of segmental bimaxillary surgery using voxel- and surface-based registration: A comparative study

Michael Boelstoft Holte [a,b], Alexandru Diaconu [a,b], Else Marie Pinholt [a,b,*]

[a] 3D Lab Denmark, Department of Oral and Maxillofacial Surgery, University Hospital of Southern Denmark, Esbjerg, Denmark
[b] Department of Regional Health Research, Faculty of Health Sciences, University of Southern Denmark, Finsensgade 35, 6700, Esbjerg, Denmark

ARTICLE INFO

Handling Editor: Dr Kaveh Shakib

Keywords:
Orthognathic surgery
Computer-assisted surgery
Three-dimensional imaging
Cone-beam computed tomography
Maxillary osteotomy
Mandibular osteotomy

ABSTRACT

The purpose of the present study was to compare the precision and reliability of voxel- and surface-based registration for computer-assisted assessment of the surgical accuracy and postoperative stability of segmental bimaxillary surgery. Three-dimensional translational and rotational measurements were performed by two observers using voxel- and surface-based registration. The precision and reliability of the measurements were calculated by the mean absolute differences (MAD) and intraclass correlation coefficients (ICC) at 95 % confidence intervals. A paired *t*-test or the non-parametric equivalent, Wilcoxon signed-rank test, was applied to statistically evaluate whether the precision of voxel- and surface-based registration was statistically significantly different ($p < 0.05$). Voxel-based registration had high precision (MAD <0.44 mm/0.92°) and excellent reliability, ICC [0.82–1.00]. The precision of surface-based registration was lower (MAD <0.56 mm/1.45°) and the reliability ranged from poor to excellent for the different bone segments, ICC [0.33–1.00]. Both registration techniques had high precision and excellent reliability for the assessment of the surgical accuracy, and the error margin of both techniques was clinical irrelevant. However, the increased precision of voxel-based registration was statistically significant ($p < 0.05$) for the maxillary segments and the chin, and the stability measurement error (ranging up to 1.58 mm and 4.46°) introduced by surface-based registration may be considered clinical relevant for these bone segments. Within the limitations of the present comparative study, voxel-based registration generally exhibited higher precision and reliability than surface-based registration for the surgical accuracy and postoperative stability assessment of segmental bimaxillary surgery.

1. Introduction

Superimposition and image registration techniques have become popular for computer-assisted assessment of craniomaxillofacial surgery and growth using cone-beam computed tomography (CBCT) [1–9]. Several recent systematic literature reviews encourage the use of registration techniques for automated and standardised three-dimensional (3D) assessment of orthognathic surgery [10–12]. Comprehensive meta-analyses of the accuracy of virtual surgical planning (VSP) or patient-specific implants in orthognathic surgery were not possible due to inhomogeneity and lack of standardization of outcome measures [12–14]. Consequently, evidence is still lacking for definitive conclusions on the performance of these applied technologies.

Gaber et al. proposed a universal protocol for the assessment of orthognathic surgery based on a six degrees-of-freedom (translational and rotational measurements in 3D) error metric between the planned and realised repositioning of the maxillofacial bone segments using image registration [10]. In extension, Andnoli et al. suggested several reference structures, which can be applied for registration in order to assess the different moving bone segments [11]. Recently, this standardisation has been accepted and adopted in the literature by numerous authors [1,2,5–9].

Voxel-based registration (VBR) aligns two CBCT images using maximization of mutual information of the voxels, and surface-based registration (SBR) aligns two surfaces, e.g. using the iterative closest point algorithm. Voxel- and surface-based registration have previously been compared for 3D assessment of orthognathic surgery, including single-piece Le Fort I osteotomy [15], orthodontic treatment [16], and for evaluation of growing subjects [17] and mandibular condyle remodelling [18]. In most of the comparative studies the two

* Corresponding author. Department of Regional Health Research, Faculty of Health Sciences, University of Southern Denmark, Finsensgade 35, 6700, Esbjerg, Denmark.

E-mail addresses: Michael.Boelstoft.Holte@rsyd.dk (M.B. Holte), Alexandru.Diaconu@rsyd.dk (A. Diaconu), empinholt@health.sdu.dk (E.M. Pinholt).

https://doi.org/10.1016/j.adoms.2023.100470
Received 3 November 2023; Accepted 4 December 2023
Available online 6 December 2023

registration techniques were proven to be accurate and reliable with statistically insignificant differences, unlikely to be of clinical significance [15–17]. However, the performance of the two registration techniques has not been compared for assessment of segmental bimaxillary surgery, such as combined multi-piece Le Fort I osteotomy, bilateral sagittal split osteotomy and genioplasty. These types of surgery are more complex and result in repositioning of several smaller bone segments. Consequently, the 3D assessment is more challenging [1]. Furthermore, the two registration techniques have not been compared for long-term (≥ two years) postoperative stability assessment. At long-term follow-up, bone changes appearance due to remodelling. This represents a challenge for the registration techniques [2]. Additionally, previous comparative studies have not been based on the broadly accepted universal protocol proposed by Gaber et al. [10].

The purpose of the present study was to compare the precision and reliability of VBR and SBR for standardised computer-assisted assessment of segmental bimaxillary surgery. The null hypothesis was: the precision and reliability of VBR and SBR are not significantly different for standardised 3D assessment of the surgical accuracy and postoperative stability of segmental bimaxillary surgery.

2. Materials and methods

The virtual surgical analysis (VSA) proposed by Holte et al. [1,2] was applied as a framework for comparison of VBR and SBR for 3D assessment of the surgical accuracy and postoperative stability of segmental bimaxillary surgery, comprising a combined three-piece Le Fort I, bilateral sagittal split and genioplasty osteotomies. The VSA was designed based on the suggested criteria by Gaber et al. [10], originally including regional VBR. An alternative version using SBR was created in order to compare the performance of the two registration techniques.

Virtual surgical analysis uses the pre- and postoperative CBCT scans and the VSP as input, and provides as output the surgical accuracy and the postoperative stability with six degrees-of-freedom; translations and rotations in the three axes (Fig. 1) [1,2]. The surgical accuracy was calculated as the translational (anterior-posterior, right-left and superior-inferior) and rotational (pitch, roll and yaw) differences between the VSP and the short-term postoperative outcome of the individual bone segments. The translational accuracy was calculated at the clinical relevant anatomical landmarks presented in Table 1 without the need of landmark re-identification in the postoperative CBCT scans [1]. Similarly, the postoperative stability was calculated as the differences between the short- and long-term postoperative outcomes (Fig. 1) [2]. The VSP was made in Dolphin Imaging® (Patterson, Chatsworth, CA). Please refer to the works by Holte et al. [1,2] for further details on study sample, image acquisition, surgical planning and intervention, and technical details on the 3D assessment.

2.1. Voxel-based registration

Voxel-based registration was performed in Mimics® 23 (Materialise NV, Leuven, Belgium) using the Automatic Image Registration tool. The individual bone segments in the preoperative CBCT scan were registered to their positions in the short- and long-term postoperative CBCT scans by VBR using the virtual bone segments of the VSP as volume of interest (VOI). For the assessment of the postoperative stability by registration on the long-term postoperative scan, the VOI were modified to exclude the teeth, which are prone to interval changes due to postsurgical orthodontics [2]. Fig. 2 shows the selected VOI for registration of the individual bone segments using VBR.

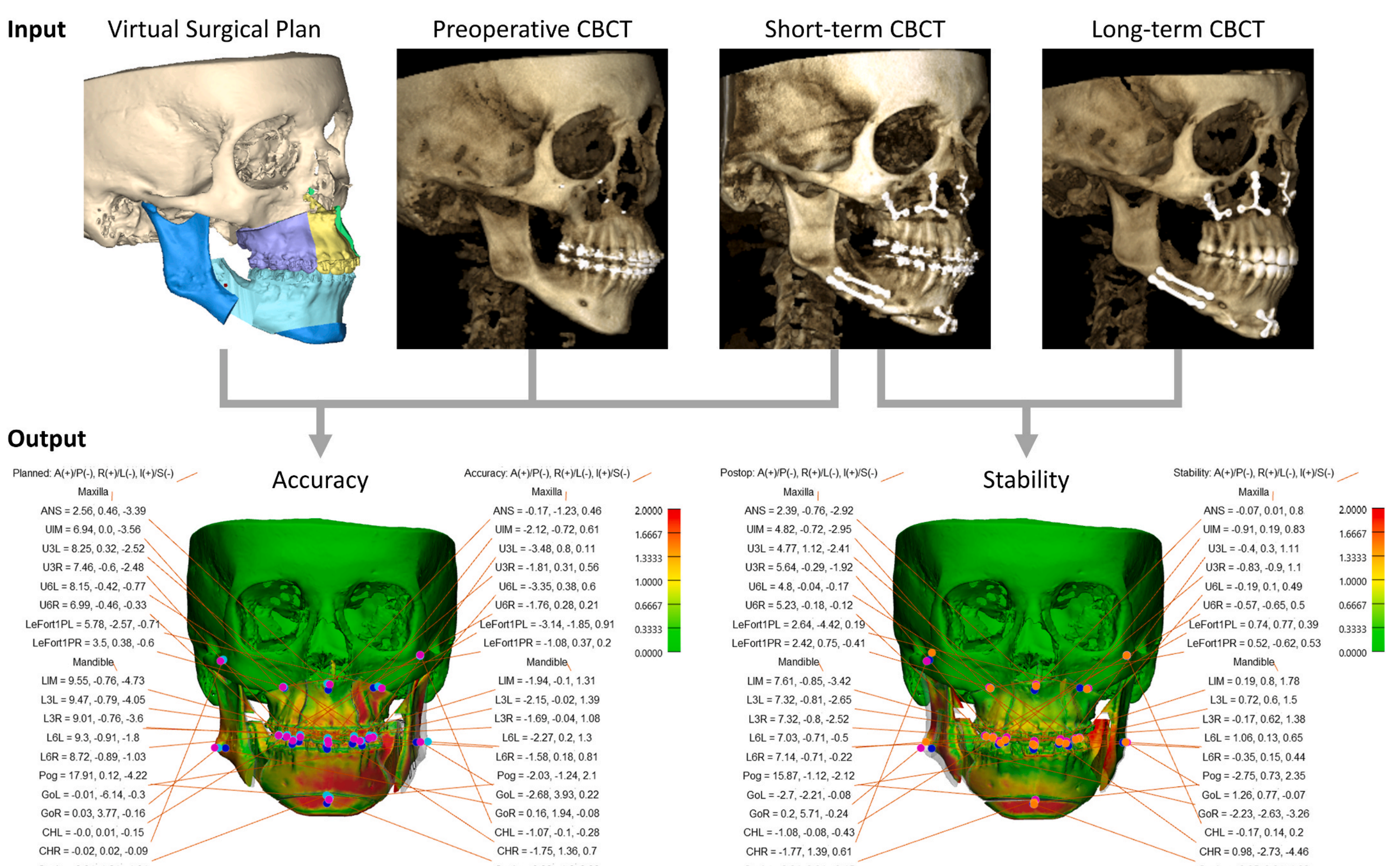

Fig. 1. Data input and output for virtual surgical analysis using voxel-and surface-based registration. Input: preoperative and short- and long-term postoperative cone-beam computed tomography (CBCT) and 3D surface models. Output: three-dimensional surgical accuracy and postoperative stability assessment.

Table 1
Definitions of 3D cephalometric landmarks.

Landmark	Definition	Bilateral
Maxillary landmarks		
Anterior nasal spine (ANS)	The most anterior midpoint of the anterior nasal spine of the maxilla.	
Upper incisors midpoint (UIM)	The most mesial point of the tip of the crown in between the upper central incisors.	
Upper canine (U3)	The most inferior point of the tip of the crown of the upper canine.	R/L
Upper molar (U6)	The most inferior point of the mesial buccal cusp of the crown of the first upper molar.	R/L
Le Fort I Osteotomy point (LF)	The most anterior point on the Le Fort I osteotomy line at the distal root of the first upper molar.	R/L
Mandibular landmarks		
Lower incisors midpoint (MID)	The most mesial point of the tip of the crown between the lower central incisors.	
Lower Canine (L3)	The most superior point of the tip of the crown of the lower canine.	R/L
Lower molar (L6)	The most superior point of the mesial buccal cusp of the crown of the lower molar.	R/L
Pogonion (Pog)	The most anterior midpoint of the chin on the outline of the mandibular symphysis.	
Condylar hinge point (CR)	The most superior posterior point on the mandibular condyle.	R/L
Gonion (Go)	The most caudal, posterior and lateral point of the mandibular angle.	R/L

R: right, L: left.

2.2. Surface-based registration

Surface-based registration was performed in 3-matic® 16 (Materialise NV, Leuven, Belgium) using the Global Registration function. Surface-based registration requires that the maxillofacial bone is segmented and 3D reconstructed from the CBCT scans for regional registration of the corresponding bone segments. The preoperative 3D models were imported from the VSP and the postoperative image segmentation was performed in Mimics® 23 (Materialise NV, Leuven, Belgium) beginning with a minimum threshold of 283 Hounsfield units (HU) followed by manual refinement.

The 3D parts were constructed with default "optimal quality" parameters: voxel resolution (x-, y- and z-multiplier: 1); no smoothing; triangle reduction (reducing mode: advanced edge, tolerance: 0.05625 mm, edge angle: 10°, iterations: 3). Fig. 2 shows the selected region of interest (ROI) for registration of the individual bimaxillary bone segments using SBR, which, although selected differently, was attempted to replicate the VOI used for VBR.

2.3. Statistical analysis

Statistical analysis of the data was performed in STATA® 16.1 (StataCorp, College Station, TX). A sample size calculation for a one-sample correlation study (corr. > 0.8, power $= 0.8$, alpha $= 0.05$, and two raters) was performed. To evaluate and compare the precision and reliability of the two registration techniques, two observers independently performed the assessment (MH, AD). After a period of minimum two weeks, one of the observers (MH) analysed the data again blinded to the initial assessment. To avoid confounding factors caused by landmark digitalisation, landmarks identified once in the preoperative CBCT scans were imported for each assessment.

The precision of the two registration methods was measured in accordance to previous studies [19], i.e. by the mean absolute difference (MAD) of the repeated measurements. A Shapiro-Wilk test ($p > 0.05$) was used to test for normality. To evaluate whether the precision of VBR and SBR was statistically significantly different ($p < 0.05$), a paired *t*-test

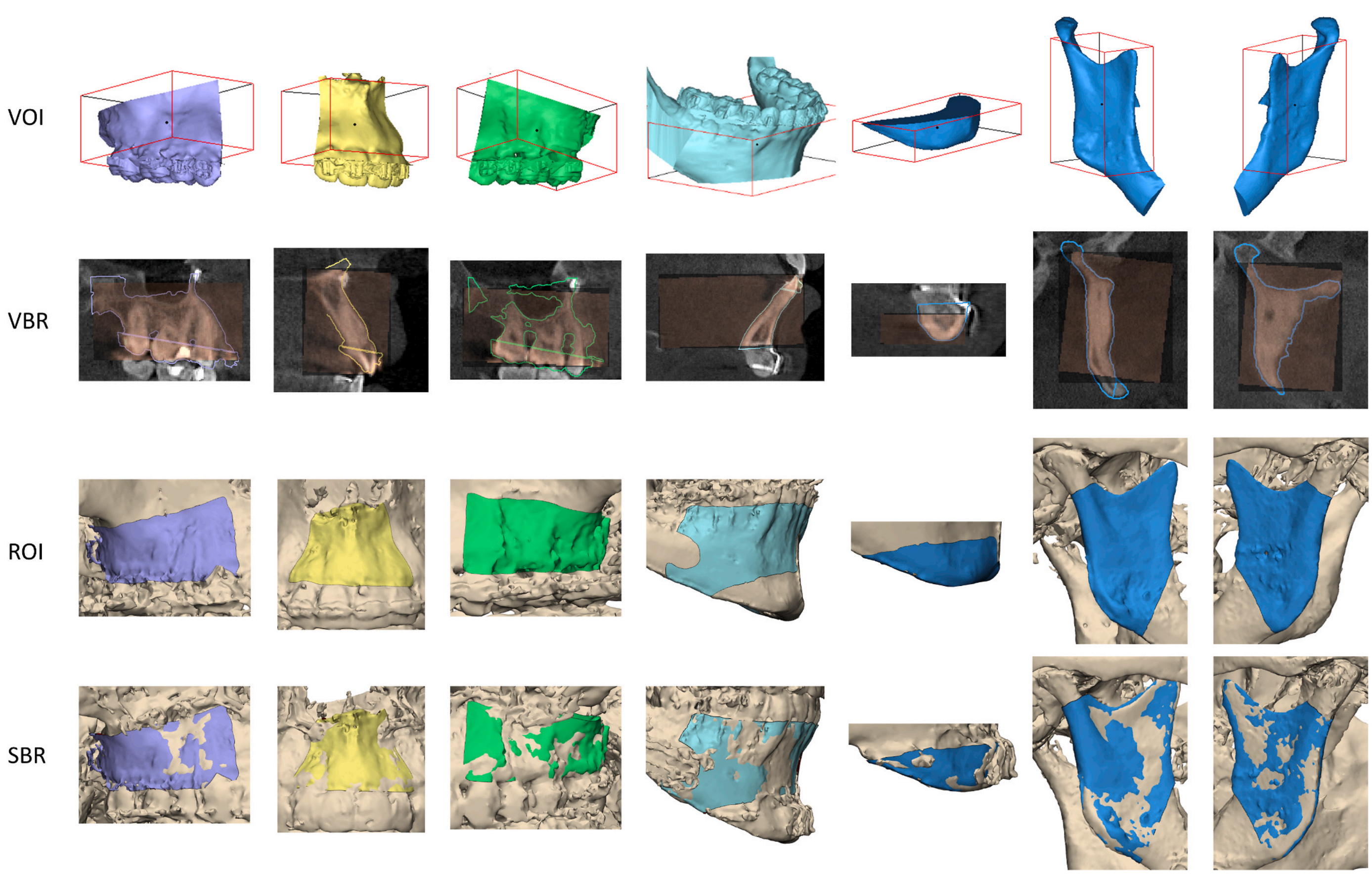

Fig. 2. Volume (VOI) and region of interest (ROI) for all repositioned bone segments used for voxel- (VBR) and surface-based registration (SBR), respectively.

was applied for normally distributed data. The non-parametric equivalent, Wilcoxon signed-rank test was applied otherwise.

In accordance to the performance evaluation criteria by Gaber et al. [10], the reliability comparison was based on the intra- and inter-observer reliability. Hence, intraclass correlation coefficients (ICC) at a 95 % confidence interval on the measurements of the two observers were calculated for single measurements using a one-way random effect model. Additionally, Bland-Altman plots were generated to evaluate observer agreement, including bias, 95 % limits of agreement and tendency of systematic error.

3. Results

The statistical sample size calculation resulted in a required sample size of n = 10 to obtain a statistical power of 0.8, which is in accordance to previous studies [1,2,5,9]. Hence, the first ten consecutive post-pubertal patients in 2019 (January–April) who met the inclusion and exclusion criteria were retrospectively included (six male; four female; mean age 24.4 years; skeletal class II and III; symmetry/asymmetry; maxillo-mandibular advancement; maxillary expansion; mandibular setback and genioplasty).

3.1. Precision

Tables 2 and 3 show the precision and reliability of the assessment of the surgical accuracy and postoperative stability using VBR and SBR. The precision of VBR was statistically significant superior for the assessment of the translational surgical accuracy of the right maxillary segment (p = 0.016) and the chin (p = 0.019), and for the translational stability of the central (p = 0.011) and left maxillary segment (p = 0.027) (Table 2). Voxel-based registration was statistically significantly more precise for the assessment of the rotational surgical accuracy of the chin (p = 0.006/0.013), while SBR was statistically significantly more precise for the assessment of the rotational surgical accuracy of the distal (p = 0.033/0.039) and right proximal mandibular segment (p = 0.001) (Table 3).

3.2. Reliability

The intra- and inter-observer reliability of the assessment of the surgical accuracy of all the bone segments using both VBR and SBR was excellent [20], ICC range [0.96–1.00] and [0.88–1.00], respectively (Tables 2 and 3). The intra- and inter-observer reliability of the assessment of the postoperative stability using VBR was excellent, ICC range [0.82–0.99]. In contrast, the ICC range was good to excellent [0.73–0.99] and poor to excellent [0.33–0.99] for the assessment of the stability using SBR, respectively. The low ICC values for SBR were observed for the stability assessment of the maxillary segments and the chin (Tables 2 and 3).

Tables 4 and 5 summarize the data of the Bland-Altman plots (available in appendix A) for further analysis of the intra- and inter-observer agreement. The observer agreement on the surgical accuracy measurements using both VBR and SBR showed low biases, narrow limits of observer agreement, and no signs of systematic errors (Table 4). The observer agreement on the stability measurements using VBR showed lower biases, and the limits of the observer agreement were also considerably narrower than for SBR (Table 5). The highest biases and largest limits of agreements were observed for the stability assessment of the maxillary segments and the chin using SBR. Again, there were no signs of systematic errors.

4. Discussion

Generally, VBR was more precise and reliable than SBR. Both registration techniques had high precision and excellent reliability for the assessment of the surgical accuracy of all repositioned bone segments, and the error margin of both techniques was clinical irrelevant. The performance difference was more pronounced for the assessment of the postoperative stability, where VBR had high precision and excellent reliability. In contrast, the precision of SBR was lower and the reliability ranged from poor to excellent for the different bone segments. The increased precision of VBR was statistically significant for the assessment of the maxillary segments and the chin, and the stability measurement error introduced by SBR for these bone segments may be clinical relevant. Consequently, the null-hypothesis was rejected.

Especially, the assessment of the postoperative stability of the smaller bone segments (the maxillary segments and the chin) using SBR showed larger mean absolute translational and rotational errors of up to 0.56 mm and 1.45°, and wide limits of agreement, ranging up to 1.58 mm and 4.46°. Such measurement errors may be of clinical relevance, according to most literature, where error margins of more than 2 mm and 4° are often consider surgical inaccurate or unstable [7,14,21–25]. Despite the lowest precision was observed for the assessment of the rotational stability using SBR, no statistically significant difference between the two registration techniques was found.

Surface-based registration demonstrated a statistically significantly

Table 2
Difference between surface- and voxel-based registration presented by intraclass correlation coefficients (ICC), mean absolute differences (MAD), standard deviations (SD) and paired *t*-test p-values of the repeated intra- and inter-observer 3D translational measurements.

Bone segment	Intra-observer agreement					Inter-observer agreement				
	Surface-based reg.		Voxel-based reg.		P	Surface-based reg.		Voxel-based reg.		P
	ICC	MAD (SD)	ICC	MAD (SD)		ICC	MAD (SD)	ICC	MAD (SD)	
Translational surgical accuracy (mm)										
Mandibular distal (LIM)	1.00	0.08 (0.07)	1.00	0.08 (0.06)	0.869	0.99	0.12 (0.11)	0.99	0.13 (0.09)	0.334
Proximal left (CHL)	0.95	0.14 (0.27)	0.99	0.08 (0.07)	0.371	0.93	0.20 (0.31)	0.99	0.12 (0.09)	0.918
Proximal right (CHR)	0.98	0.10 (0.08)	0.99	0.07 (0.05)	0.125	0.98	0.11 (0.11)	0.96	0.17 (0.15)	0.082
Chin (Pog)	0.96	0.21 (0.24)	0.99	0.11 (0.13)	0.019[a]	0.88	0.27 (0.47)	0.98	0.19 (0.18)	0.926
Maxillary central (UIM)	0.98	0.15 (0.17)	0.99	0.12 (0.09)	0.845	0.95	0.25 (0.32)	0.98	0.20 (0.19)	0.797
Maxillary left (U3L)	0.98	0.17 (0.18)	0.99	0.10 (0.08)	0.074	0.97	0.19 (0.17)	0.99	0.14 (0.11)	0.092
Maxillary right (U3R)	0.99	0.14 (0.12)	0.99	0.10 (0.10)	0.125	0.99	0.17 (0.14)	1.00	0.09 (0.08)	0.016[a]
Translational postoperative stability (mm)										
Mandibular distal (LIM)	0.98	0.16 (0.22)	0.99	0.14 (0.13)	0.711	0.99	0.17 (0.16)	0.97	0.28 (0.27)	0.125
Proximal left (CHL)	0.94	0.26 (0.26)	0.98	0.18 (0.17)	0.309	0.94	0.27 (0.24)	0.93	0.26 (0.36)	0.558
Proximal right (CHR)	0.99	0.19 (0.18)	0.99	0.14 (0.13)	0.181	0.99	0.23 (0.19)	0.94	0.44 (0.50)	0.150
Chin (Pog)	0.90	0.25 (0.23)	0.95	0.23 (0.21)	0.975	0.79	0.38 (0.43)	0.84	0.43 (0.38)	0.253
Maxillary central (UIM)	0.91	0.38 (0.50)	0.97	0.28 (0.25)	0.781	0.86	0.56 (0.46)	0.97	0.30 (0.22)	0.011[a]
Maxillary left (U3L)	0.84	0.40 (0.39)	0.96	0.23 (0.21)	0.027[a]	0.73	0.48 (0.53)	0.94	0.28 (0.24)	0.070
Maxillary right (U3R)	0.94	0.25 (0.20)	0.96	0.23 (0.21)	0.658	0.87	0.35 (0.39)	0.96	0.25 (0.22)	0.285

[a] Statistically significant.

Table 3
Difference between surface- and voxel-based registration presented by intraclass correlation coefficients (ICC), mean absolute differences (MAD), standard deviations (SD) and paired *t*-test p-values of the repeated intra- and inter-observer 3D rotational measurements.

Bone segment	Intra-observer agreement					Inter-observer agreement				
	Surface-based reg.		Voxel-based reg.		P	Surface-based reg.		Voxel-based reg.		P
	ICC	MAD (SD)	ICC	MAD (SD)		ICC	MAD (SD)	ICC	MAD (SD)	
Rotational surgical accuracy (°)										
Mandibular distal (LIM)	1.00	0.07 (0.08)	0.99	0.11 (0.08)	0.033[a]	0.98	0.12 (0.19)	0.98	0.17 (0.17)	0.039[a]
Proximal left (CHL)	1.00	0.13 (0.16)	1.00	0.12 (0.09)	0.681	1.00	0.13 (0.17)	1.00	0.16 (0.08)	0.092
Proximal right (CHR)	1.00	0.12 (0.18)	1.00	0.12 (0.10)	0.517	1.00	0.10 (0.12)	1.00	0.24 (0.23)	0.001[a]
Chin (Pog)	0.93	0.70 (0.74)	0.99	0.33 (0.26)	0.006[a]	0.90	0.84 (1.04)	0.96	0.52 (0.65)	0.013[a]
Maxillary central (UIM)	0.99	0.29 (0.42)	0.99	0.35 (0.32)	0.217	0.97	0.44 (0.68)	0.98	0.63 (0.42)	0.055
Maxillary left (U3L)	0.98	0.35 (0.44)	0.99	0.28 (0.21)	0.789	0.98	0.43 (0.39)	0.99	0.37 (0.30)	0.472
Maxillary right (U3R)	1.00	0.20 (0.23)	1.00	0.21 (0.18)	0.319	0.99	0.30 (0.34)	1.00	0.22 (0.15)	0.371
Rotational postoperative stability (°)										
Mandibular distal (LIM)	0.97	0.25 (0.38)	0.99	0.20 (0.16)	0.622	0.97	0.26 (0.35)	0.94	0.33 (0.43)	0.558
Proximal left (CHL)	0.98	0.25 (0.30)	0.98	0.21 (0.23)	0.837	0.99	0.21 (0.19)	0.99	0.23 (0.18)	0.688
Proximal right (CHR)	0.97	0.29 (0.32)	0.99	0.22 (0.22)	0.074	0.98	0.25 (0.26)	0.94	0.47 (0.49)	0.098
Chin (Pog)	0.73	0.87 (0.74)	0.91	0.59 (0.51)	0.082	0.67	1.26 (1.05)	0.82	0.92 (0.78)	0.192
Maxillary central (UIM)	0.84	0.86 (0.93)	0.95	0.63 (0.63)	0.537	0.47	1.45 (1.56)	0.93	0.88 (0.66)	0.206
Maxillary left (U3L)	0.59	0.91 (1.19)	0.92	0.59 (0.67)	0.504	0.33	1.14 (1.54)	0.90	0.78 (0.58)	0.382
Maxillary right (U3R)	0.94	0.46 (0.37)	0.95	0.48 (0.54)	0.829	0.86	0.78 (0.64)	0.94	0.60 (0.45)	0.241

[a] Statistically significant.

Table 4
Summary of Bland-Altman plots of intra- and inter-observer agreement with bias, upper (UL) and lower limits (LL) of agreement for translational measurements.

Bone segment	Intra-observer agreement						Inter-observer agreement					
	Surface-based reg.			Voxel-based reg.			Surface-based reg.			Voxel-based reg.		
	Bias	LL	UL	Bias	LL	UL	Bias	LL	UL	Bias	LL	UL
Translational surgical accuracy (mm)												
Mandibular distal (LIM)	0.02	−0.18	0.22	0.01	−0.19	0.21	0.02	−0.30	0.34	0.03	−0.28	0.34
Proximal left (CHL)	0.09	−0.49	0.67	0.00	−0.22	0.22	0.06	−0.65	0.78	0.01	−0.29	0.31
Proximal right (CHR)	0.00	−0.26	0.26	0.00	−0.18	0.18	0.03	−0.27	0.33	−0.01	−0.46	0.45
Chin (Pog)	0.04	−0.58	0.66	−0.02	−0.34	0.31	−0.07	−1.13	1.00	−0.11	−0.57	0.34
Maxillary central (UIM)	−0.05	−0.48	0.39	0.00	−0.30	0.30	0.10	−0.66	0.87	−0.01	−0.55	0.53
Maxillary left (U3L)	0.03	−0.46	0.52	−0.01	−0.27	0.25	0.04	−0.45	0.54	0.02	−0.34	0.37
Maxillary right (U3R)	0.02	−0.34	0.38	0.00	−0.28	0.28	−0.04	−0.46	0.39	−0.04	−0.27	0.19
Translational postoperative stability (mm)												
Mandibular distal (LIM)	0.01	−0.54	0.55	0.03	−0.34	0.41	0.05	−0.39	0.50	−0.19	−0.85	0.46
Proximal left (CHL)	−0.09	−0.79	0.62	−0.01	−0.50	0.48	−0.13	−0.80	0.54	−0.07	−0.94	0.80
Proximal right (CHR)	−0.03	−0.54	0.48	0.00	−0.37	0.36	−0.07	−0.64	0.51	−0.15	−1.43	1.13
Chin (Pog)	0.00	−0.67	0.66	0.06	−0.55	0.67	0.11	−0.99	1.22	0.09	−1.03	1.21
Maxillary central (UIM)	0.05	−1.19	1.28	0.01	−0.73	0.75	0.21	−1.17	1.58	−0.02	−0.76	0.72
Maxillary left (U3L)	−0.12	−1.20	0.97	−0.02	−0.63	0.60	−0.12	−1.51	1.26	0.04	−0.69	0.76
Maxillary right (U3R)	−0.10	−0.70	0.50	0.01	−0.62	0.63	−0.09	−1.11	0.92	−0.08	−0.72	0.56

Table 5
Summary of Bland-Altman plots of intra- and inter-observer agreement with bias, upper (UL) and lower limits (LL) of agreement for rotational measurements.

Bone segment	Intra-observer agreement						Inter-observer agreement					
	Surface-based reg.			Voxel-based reg.			Surface-based reg.			Voxel-based reg.		
	Bias	LL	UL	Bias	LL	UL	Bias	LL	UL	Bias	LL	UL
Rotational surgical accuracy (°)												
Mandibular distal (LIM)	0.00	−0.21	0.21	0.01	−0.26	0.28	0.03	−0.40	0.47	0.01	−0.46	0.49
Proximal left (CHL)	−0.03	−0.44	0.37	−0.01	−0.30	0.29	0.03	−0.39	0.44	0.02	−0.32	0.37
Proximal right (CHR)	0.06	−0.34	0.47	−0.02	−0.32	0.28	0.08	−0.18	0.35	0.08	−0.56	0.72
Chin (Pog)	0.16	−1.82	2.15	0.05	−0.76	0.87	0.16	−2.46	2.79	−0.03	−1.68	1.62
Maxillary central (UIM)	0.01	−1.00	1.02	−0.06	−0.99	0.86	0.01	−1.58	1.59	0.01	−1.50	1.51
Maxillary left (U3L)	0.03	−1.08	1.13	0.12	−0.53	0.77	0.11	−1.01	1.24	0.23	−0.60	1.06
Maxillary right (U3R)	−0.05	−0.64	0.55	0.01	−0.54	0.56	−0.12	−1.00	0.75	−0.08	−0.57	0.41
Rotational postoperative stability (°)												
Mandibular distal (LIM)	0.13	−0.73	0.99	0.00	−0.51	0.51	0.07	−0.77	0.92	−0.09	−1.14	0.96
Proximal left (CHL)	0.08	−0.67	0.83	0.00	−0.60	0.61	0.04	−0.52	0.60	0.02	−0.56	0.60
Proximal right (CHR)	0.22	−0.52	0.95	−0.03	−0.63	0.58	0.06	−0.63	0.76	−0.11	−1.43	1.21
Chin (Pog)	−0.21	−2.44	2.02	−0.14	−1.66	1.38	−0.22	−3.44	2.99	0.08	−2.29	2.46
Maxillary central (UIM)	−0.37	−2.76	2.02	−0.17	−1.90	1.56	−0.30	−4.46	3.85	0.02	−2.16	2.21
Maxillary left (U3L)	−0.44	−3.26	2.37	−0.12	−1.86	1.61	−0.61	−4.20	2.98	−0.05	−1.98	1.89
Maxillary right (U3R)	−0.07	−1.23	1.09	0.05	−1.37	1.47	−0.08	−2.07	1.92	0.05	−1.44	1.53

higher precision for the assessment of the rotational surgical accuracy of the distal and right proximal mandibular segments. However, in these cases both registration techniques had a very high precision, excellent reliability and a clinically irrelevant error margin.

The present study found statistically significant differences between VBR and SBR. Previous comparative studies also found that SBR had higher variability than VBR [15,16]. However, the difference were statistically insignificant. One of these studies also concluded that SBR has lower efficiency, as it requires more manual input in form of segmentation and 3D reconstruction of the surface models [16]. Image segmentation and surface generation is time consuming and prone to observer variability [15,26]. Voxel-based registration directly matches the CBCT images using the full intensity information of the voxels, without the need of image segmentation and surface generation. However, recently, automated image segmentation of the craniomaxillofacial bone using artificial intelligence has been proposed and validated, eliminating the need of manual image segmentation, and thereby eliminating the introduction of additional observer variability using SBR [27]. It should be noted that the Mimics and 3-matic (Materialise NV, Leuven, Belgium) functions used in the present study to perform VBR and SBR require manual input by the operator in form of pre-alignment and selection of VOI/ROI for each bone segment, which cause observer variability.

A synthesis of the findings of previous studies suggests that VBR seem to be sensitive to registration errors when using smaller unstable reference structures as VOI. This was concluded both in the study by Holte et al. [18], who used a part of the ramus, which was proven to remodel after surgery, for the assessment of condylar remodelling; and in the study by Ruellas et al., who used the small reference structure proposed by Björk et al. to evaluating growing subjects, and resulted in miss-registration in 14 out of 16 cases [28]. To encounter poor- or miss-registration, the present study applied the entire bone segments from the VSP as VOI for regional registration. However, for assessment of the postoperative stability, the regions was modified not to include the teeth, which are prone to interval changes due to postsurgical orthodontics (Fig. 2).

The contribution of the present comparative study is threefold. To the best of the authors' knowledge, (1) this is the first study to compare the performance of VBR and SBR for the assessment of orthognathic surgery based on the broadly accepted and adopted universal protocol proposed by Gaber et al. [10] Previous comparative studies quantified the outcome inhomogeneously or using less clinical relevant measurements, e.g. surface distances, representing an average discrepancy of the surface of the bone segment. Such measures are different from the translational and rotational repositioning, which are used in the VSP of orthognathic surgery, and hence, difficult to use for clinical evaluation of the surgical outcome. (2) The present study is the first study to compare the performance of VBR and SBR for the assessment of the outcome of more complex segmental bimaxillary surgical procedures, and (3) for the assessment of the postoperative stability of orthognathic surgery.

The limitations of the present comparative study were as follows. The CBCT scans were produced by a single device - Planmeca Viso™ G7 scanner (Planmeca OY, Helsinki, Finland), and VBR and SBR was performed in one, however, frequently used commercial software package - Mimics Innovation Suite (Materialise NV, Leuven, Belgium). Additionally, the image segmentation and the parameters of the 3D surface construction may influence SBR.

5. Conclusions

Within the limitations of the present comparative study, VBR generally exhibited higher precision and reliability than SBR for the 3D assessment of the surgical accuracy and postoperative stability of segmental bimaxillary surgery. Both registration techniques had high precision and excellent reliability for the assessment of the surgical accuracy, and the error margin of both techniques was clinical irrelevant. For the assessment of the postoperative stability, only VBR had high precision and excellent reliability. In contrast, the precision of SBR was lower and the reliability ranged from poor to excellent for the different bone segments. The increased precision of VBR was statistically significant for the assessment of the maxillary segments and the chin, and the stability measurement error introduced by SBR for these bone segments may be clinical relevant.

Ethics approval

Permission was granted by the Institutional Ethics Committee – University Hospital of Southern Denmark, Esbjerg (21/33871). As the study comprised retrospective material, none of the subjects were exposed to either additional radiation or examinations.

Patient permission/consent

Not required.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors thank senior surgeons Janne Ingerslev and Jens Jørgen Thorn, Department of Oral and Maxillofacial Surgery, University Hospital of Southern Denmark, for clinical counsel; and Professor Gabriele Berg-Beckhoff, Department of Public Health, Faculty of Health Sciences, University of Southern Denmark, for statistical counsel.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.adoms.2023.100470.

References

[1] Holte MB, Diaconu A, Ingerslev J, et al. Virtual analysis of segmental bimaxillary surgery: a validation study. J Oral Maxillofac Surg 2021;79:2320–33.
[2] Holte MB, Diaconu A, Ingerslev J, et al. Virtual surgical analysis: long-term cone beam computed tomography stability assessment of segmental bimaxillary surgery. Int J Oral Maxillofac Surg 2022;51:1188–96.
[3] Diaconu A, Holte MB, Cattaneo PM, et al. A semi-automatic approach for longitudinal 3D upper airway analysis using voxel-based registration. Dentomaxillofac Radiol 2022;51:20210253.
[4] Holte MB, Saederup H, Pinholt EM. A semi-automatic approach for Holistic 3D assessment of Temporomandibular Joint changes. J Pers Med 2023:13.
[5] Baan F, Liebregts J, Xi T, et al. A new 3D tool for assessing the accuracy of bimaxillary surgery: the OrthoGnathicAnalyser. PLoS One 2016;11:e0149625.
[6] Baan F, Sabelis JF, Schreurs R, et al. Validation of the OrthoGnathicAnalyser 2.0-3D accuracy assessment tool for bimaxillary surgery and genioplasty. PLoS One 2021;16:e0246196.
[7] Shaheen E, Shujaat S, Saeed T, et al. Three-dimensional planning accuracy and follow-up protocol in orthognathic surgery: a validation study. Int J Oral Maxillofac Surg 2019;48:71–6.
[8] da Costa Senior O, Vaes L, Mulier D, et al. Three dimensional assessment of segmented Le Fort I osteotomy planning and follow-up: a validation study. J Dent 2021;111:103707.
[9] Stokbro K, Thygesen T. A 3-dimensional approach for analysis in orthognathic surgery-using free software for voxel-based alignment and Semiautomatic measurement. J Oral Maxillofac Surg 2018;76:1316–26.

[10] Gaber RM, Shaheen E, Falter B, et al. A systematic review to Uncover a universal protocol for accuracy assessment of 3-dimensional virtually planned orthognathic surgery. J Oral Maxillofac Surg 2017;75:2430–40.
[11] Andriola FO, Haas Junior OL, Guijarro-Martinez R, et al. Computed tomography imaging superimposition protocols to assess outcomes in orthognathic surgery: a systematic review with comprehensive recommendations. Dentomaxillofac Radiol 2022;51:20210340.
[12] Diaconu A, Holte MB, Berg-Beckhoff G, et al. Three-dimensional accuracy and stability of personalized implants in orthognathic surgery: a systematic review and a meta-analysis. J Pers Med 2023:13.
[13] Alkhayer A, Piffko J, Lippold C, et al. Accuracy of virtual planning in orthognathic surgery: a systematic review. Head Face Med 2020;16:34.
[14] Tondin GM, Leal M, Costa ST, et al. Evaluation of the accuracy of virtual planning in bimaxillary orthognathic surgery: a systematic review. Br J Oral Maxillofac Surg 2022;60:412–21.
[15] Almukhtar A, Ju X, Khambay B, et al. Comparison of the accuracy of voxel based registration and surface based registration for 3D assessment of surgical change following orthognathic surgery. PLoS One 2014;9:e93402.
[16] Han G, Li J, Wang S, et al. A comparison of voxel- and surface-based cone-beam computed tomography mandibular superimposition in adult orthodontic patients. J Int Med Res 2021;49:300060520982708.
[17] Ghoneima A, Cho H, Farouk K, et al. Accuracy and reliability of landmark-based, surface-based and voxel-based 3D cone-beam computed tomography superimposition methods. Orthod Craniofac Res 2017;20:227–36.
[18] Holte MB, Saederup H, Pinholt EM. Comparison of surface- and voxel-based registration on the mandibular ramus for long-term three-dimensional assessment of condylar remodelling following orthognathic surgery. Dentomaxillofac Radiol 2022;51:20210499.
[19] Bazina M, Cevidanes L, Ruellas A, et al. Precision and reliability of Dolphin 3-dimensional voxel-based superimposition. Am J Orthod Dentofacial Orthop 2018;153:599–606.
[20] Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol Assess 1994;6:284.
[21] Hsu SS, Gateno J, Bell RB, et al. Accuracy of a computer-aided surgical simulation protocol for orthognathic surgery: a prospective multicenter study. J Oral Maxillofac Surg 2013;71:128–42.
[22] Li B, Shen S, Jiang W, et al. A new approach of splint-less orthognathic surgery using a personalized orthognathic surgical guide system: a preliminary study. Int J Oral Maxillofac Surg 2017;46:1298–305.
[23] Tonin RH, Iwaki Filho L, Yamashita AL, et al. Accuracy of 3D virtual surgical planning for maxillary positioning and orientation in orthognathic surgery. Orthod Craniofac Res 2020;23:229–36.
[24] Tran NH, Tantidhnazet S, Raocharernporn S, et al. Accuracy of three-dimensional planning in surgery-first orthognathic surgery: planning versus outcome. J Clin Med Res 2018;10:429–36.
[25] Zhang N, Liu S, Hu Z, et al. Accuracy of virtual surgical planning in two-jaw orthognathic surgery: comparison of planned and actual results. Oral Surg Oral Med Oral Pathol Oral Radiol 2016;122:143–51.
[26] Koerich L, Weissheimer A, de Menezes LM, et al. Rapid 3D mandibular superimposition for growing patients. Angle Orthod 2017;87:473–9.
[27] Gillot M, Baquero B, Le C, et al. Automatic multi-anatomical skull structure segmentation of cone-beam computed tomography scans using 3D UNETR. PLoS One 2022;17:e0275033.
[28] Ruellas AC, Yatabe MS, Souki BQ, et al. 3D mandibular superimposition: comparison of regions of reference for voxel-based registration. PLoS One 2016;11:e0157625.