Available online at www.sciencedirect.com
ScienceDirect
British Journal of Oral and Maxillofacial Surgery 60 (2022) 412–421
BRITISH Journal of Oral and Maxillofacial Surgery
www.bjoms.com

Review

Evaluation of the accuracy of virtual planning in bimaxillary orthognathic surgery: a systematic review

Gustavo Marques Tondin [a], Marilia de Oliveira Coelho Dutra Leal [b], Sarah Teixeira Costa [c,*], Ricardo Grillo [a], Claudio Roberto Pacheco Jodas [a], Rubens Gonçalves Teixeira [a]

[a] *São Leopoldo Mandic College, Buccomaxillofacial Surgery and Traumatology Department, Campinas, São Paulo, Brazil*
[b] *Institute of Legal Medicine – RR, Brazil/Pos PhD Pathology – Histology Department, Dentistry School, Campinas State University, Piracicaba, São Paulo, Brazil*
[c] *Institute of Criminalistics– SP, Brazil/Dentistry School, Presidente Tancredo Neves University Center, São João del Rei, Minas Gerais, Brazil*

Received 25 May 2021; accepted in revised form 5 September 2021
Available online 20 September 2021

Abstract

The purpose of this research was to evaluate the accuracy of virtual planning in bimaxillary orthognathic surgery in bone by comparing the mean linear and angular measurements of the surgical plan with the actual surgical result. Electronic databases, MEDLINE via PubMed, Web of Science, SCOPUS, the Cochrane Library, grey literature, and the American clinical trials registry (www.ClinicalTrials.gov), were accessed as search engines. The studies consisted of publications on the assessment of accuracy in virtual planning in bimaxillary orthognathic surgery between 2010 and 2020. After application of the eligibility criteria, 26 articles were included, and their quality was evaluated using the methodological index for non-randomised studies (MINORS) tool and Cohen's kappa statistic in the MedCalc program (MedCalc Software Ltd). Evidence obtained by comparing the planning and surgical results, both in the maxilla and mandible, showed that there is great accuracy in virtual planning in bimaxillary orthognathic surgery.

Keywords: Bimaxillary orthognathic surgery; Virtual surgical planning; Accuracy

Introduction

Patients with dentofacial deformities often require surgical correction to improve functional limitations, especially in chewing and speaking, and to achieve occlusal and facial harmony. These deformities can be associated with one or two bony bases, and involve the three axes (vertical, horizontal, and transverse). Severe dental malocclusion and severe skeletal changes can require a combined approach to orthodontics and orthognathic surgery, a treatment choice that has been widely used to correct dentofacial deformities in the past three decades.[1]

Conventional methods can be used to plan orthognathic surgery, assembling plaster models on semi-adjustable articulators, and performing cephalometric analysis on teleradiography. Later, surgery on plaster models is done and a new maxillomandibular-relation surgical guide produced. However, the use of traditional guides often results in inaccuracies. Articulators also cause inaccuracies, as the rotational axis may differ from the patient's bicondylar axis, and the horizontal reference plane may differ from the registered Frankfurt plane.[2] Facebow registration to transfer the patient's condylar/dental relation can be altered by approaching the Frankfurt plane, and it can be difficult to position the facial arch on the soft tissues. The registered position of the mandible can also be modified during transposition to a semi-adjustable articulator. Plaster models and radiographs may be distorted and may present some limitations. Thus, the accuracy of measurement is impaired, especially in cases involving asymmetries of the frontal plane.[3,4]

With the development of 3-dimensional (3D) computerised technology, the planning of orthognathic surgery has changed from conventional clinical assessment with cephalometric planning and plaster model surgery to the use of

https://doi.org/10.1016/j.bjoms.2021.09.010
0266-4356/

* Corresponding author at: 120, Conselheiro Rodrigues Alves Square, Guaratinguetá, São Paulo CEP: 12501-230, Brazil.
E-mail address: sarahteixeiracosta@yahoo.com.br (S. T. Costa).

computer-aided design/computer-aided manufacturing (CAD/CAM) and virtual surgical planning (VSP). Three-dimensional models of the skull from computed tomography (CT) assist in analysis, diagnosis, and surgical planning.[1,5]

Three-dimensional image reconstruction from CT offers the possibility to perform virtual surgeries and to customise surgical guides (splints) using CAD/CAM software technology. Tooth records, cephalometric points, voxels, lines, and angles, which can be obtained using software tools, make it possible to compare the virtual plan with the postoperative surgical result.[6] Currently, three registration methods (registration based on points, surface, or voxel) can be used to superimpose postoperative 3D images on the images planned preoperatively.[7,8]

The main objective of the present study was to evaluate the accuracy of virtual planning in bimaxillary orthognathic surgery in bone through a systematic review. Specific aims were to review and critically evaluate the methods used for virtually planned orthognathic surgery regarding image, software, registration method, success criteria, and accuracy values; also to verify the need for the use of printed 3D surgical guides in virtual planning, and to find or suggest a standard protocol for virtual planning in bimaxillary orthognathic surgery.

Material and methods

The protocol was registered with the International Prospective Register of Systematic Reviews (PROSPERO) under number CRD42020183168.

Search strategy

This systematic review included currently available literature written only in English. The search used MEDLINE via PubMed, Web of Science, SCOPUS, the Cochrane Library database, grey literature, and the record of American clinical studies (www.ClinicalTrials.gov). The grey literature search was performed using Google Scholar, Open Grey, and ProQuest Dissertations and Theses databases. The final search was carried out on 3 August 2020. References cited in the included studies were also checked.

The inclusion criteria comprised randomised or non-randomised, prospective and retrospective studies, published between 2010 and 2020. Study samples should contain adults with dentofacial deformities who underwent bimaxillary orthognathic surgery that included virtual planning and acquisition of a surgical guide, followed by comparison of the accuracy of the surgical result (maxilla and/or mandible). Isolated maxillary or mandibular surgeries, publications with fewer than five patients, and studies that included patients with a history of previous orthognathic surgery were excluded. Those that included combined surgeries with temporomandibular joint (TMJ) interventions, or patients who had cleft palate and/or facial trauma or tumour sequelae, were also excluded.

The search strategy used controlled and uncontrolled terms with Boolean operators selected and adapted for searches in each database (Table 1 online only). Descriptors such as “dentofacial deformities”, “orthognathic surgical procedures”, “computer-aided design”, “three-dimensional imaging”, “dimensional measurement accuracy”, “virtual planning”, “bimaxillary”, “CAD-CAM”, “computer-assisted”, were used. All references were managed with specialised software (EndNote X9, Thomson Reuters), and duplicates were removed.

Study selection

The study selection was performed in two phases. In the first, two researchers independently reviewed the titles and abstracts of all the references obtained from the databases. References that did not fulfill the inclusion criteria were excluded. In the second phase, the same authors applied the inclusion criteria when reading the full texts. The lists of references in the selected studies were also critically assessed by both examiners. Any disagreement that occurred in the first or second phase was solved through discussion and mutual agreement between the authors. In situations of disagreement and/or discrepancy, a third author was called upon to reach the final decision.

Data extraction

Two authors independently collected data from the selected studies. In all cases the following variables were collected: authors, year, country of publication, type of study, sample size, average age (range), distribution by sex, radiographic method, software used for planning, method of registration, method of precision planning, success criteria, accuracy, and main conclusion.

Two authors independently assessed the quality (risk of bias) of each article using the methodological index for non-randomised studies (MINORS). Bias risks were scored as ‘0’ (not reported), ‘1’ (inappropriately reported) or ‘2’ (reported appropriately). Ideal final scores would be 16 points for non-comparative studies and 24 for comparative studies. Disagreements between the researchers were resolved through a third author. For study selection, screening of titles and abstracts, eligibility after reading the full text, as well as for assessing the risk of bias in the included studies, kappa (k) indexes were obtained using MedCalc Statistical Software version 14.8 .1 (MedCalc Software Ltd).

Comparison between the 3D plan and the surgical result (measured by linear and/or angular differences) was done to evaluate the accuracy of virtual planning in the selected studies.

Results

In the first phase of the study selection, 3890 articles were obtained from the four databases. After removing duplicates, 81 of the 2194 remaining studies were selected for the second

Table 1
Searches and MeSH terms.

Medline (PubMed)	((Dentofacial Deformities[Mesh Terms]) OR (Orthognathic Surgical procedures[MeSH Terms]) OR (Computer-Aided Design, Three-Dimensional Imaging[MeSH Terms]) OR (Dimensional Measurement Accuracy[Mesh Terms])) AND ((class II[tw] OR class III[tw] OR orthognathic[tw] OR maxill*[tw] OR craniofacial[tw] OR bimax*[tw] OR virtual planning[tw] OR Computer-Assisted [tw] OR Computer-Guided[tw] OR CAD-CAM[tw]) AND (accur*[tw] OR asses*[tw] OR precis*[tw] OR valid*[tw] OR reliabil* [tw]))
Web of Science	(Ts=("Dentofacial Deformities" OR "Orthognathic Surgical procedures" OR "Computer-Aided Design" OR "Three-Dimensional Imaging" OR "Dimensional Measurement Accuracy") AND ts=("Class II" OR "Class III" OR "orthognathic" OR "maxill" OR "craniofacial" OR "bimax*" OR "virtual planning" OR "Computer-Assisted" OR "Computer-Guided" OR "CAD-CAM") AND ts= ("accur*" OR "asses*" OR "precis*" OR "valid*" OR "reliabil*")) Índices=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, ESCI
Scopus	TITLE-ABS-KEY(("Dentofacial Deformities") OR ("Orthognathic Surgical procedures") OR ("Computer-Aided Design, Three-Dimensional Imaging") OR ("Dimensional Measurement Accuracy") AND ("Class II" OR "Class III" OR "orthognathic" OR "maxill" OR "craniofacial" OR "bimax*" OR "virtual planning" OR "Computer-Assisted" OR "Computer-Guided" OR "CAD-CAM") AND ("accur*" OR "asses*" OR "precis*" OR "valid*" OR "reliabil*"))
Cochrane Library	((Dentofacial Deformities [MeSH]) OR (Orthognathic Surgical procedures[MeSH]) OR (Computer-Aided Design, Three-Dimensional Imaging[MeSH]) OR (Dimensional Measurement Accuracy[MeSH])) AND ((class II OR class III OR orthognathic OR maxill* OR craniofacial OR bimax* OR virtual planning OR Computer-Assisted OR Computer-Guided OR CAD-CAM)) AND (accur* OR asses* OR precis* OR valid* OR reliabil*))

phase. No reference was identified from the grey literature search, and only two articles were identified from the reference list searches. In the next phase, after reading the 81 full-text articles selected, 55 were excluded. Therefore, only 26 met the eligibility criteria and were included in the qualitative analysis (Fig. 1). The characteristics of the included studies are summarised in Tables 2 and 3 (both online only) [3–5,7,9–30] (considering the directions as 'x' = medium lateral (sagittal), 'y' = anteroposterior (horizontal), 'Z' = upper-lower (vertical); and 'pitch', 'roll' and 'yaw' as rotational). The risk of bias of the included studies is identified in Table 4 (online only), with an index of agreement between reviewers (kappa) of 0.86.

Discussion

This systematic review was developed to assess the accuracy of virtual planning using 3D surgical guides in bimaxillary orthognathic surgery by comparing VSP and actual surgical outcome. Of the 26 studies, 14 evaluated the accuracy of VSP with the surgical result prospectively, while 12 obtained data retrospectively. Regarding the number of samples, the populations in the studies ranged from 6 [26,28,30] to 100 patients.[5]

Conventional treatment of dentofacial discrepancy usually involves planning based on three types of 2-dimensional images (2D) (panoramic, teleradiographic, and posteroanterior), paper planning, and surgery on plaster models. However, these can provide insufficient information about the actual movement of 3D structures during surgery and, despite careful simulation and prediction the surgical result may still be different from the original plan. The 3D planning of surgical movements is therefore essential to achieve greater accuracy.[3]

Four studies compared the accuracy of virtual and conventional 2D planning. Schneider et al[13] reported that VSP contributed to a reduction in surgical time and an improvement in accuracy. Bengtsson et al[4] showed equal accuracy for both methods (2D and 3D) studied, but the 3D technique had an obvious advantage in patients with facial asymmetry. Ritto et al[18] reported that orthognathic surgery is a precise technique and VSP presents accuracy similar to planning with a semi-adjustable articulator. Zinser and Zoeller[23] found greater accuracy in VSP when compared with a conventional 2D technique.

Different methodologies have been applied to measure discrepancies between VSP and the postoperative results. However, a recent study showed a lack of consensus in the literature regarding the methods used to evaluate accuracy in virtually planned orthognathic surgery.[6]

VSP and 3D printing of surgical guides are becoming standard for the correction of dentofacial deformities.[7,22] In simulation software, VSP is transferred to the patient using guides, which can be manufactured directly and printed using CAD/CAM technology.[7]

Three methods can be used for 3D overlap in clinical diagnosis and evaluation of treatment results: 'voxel'-based; reference point-based; and surface-based. [7,8] It is important to note that 3D overlay remains challenging and is much more complicated than 2D overlay. Difficulties in assessing the reliability of 3D overlays not only reflected current problems, but were also the result of the choices of regions used to test the reproducibility of the overlay (points on various surfaces in the three spatial planes).[8]

In the voxel-based registration method, all the steps are automated to ensure that analysis is independent of observer error. The method has been widely described in the literature to assess changes after orthognathic surgery and orthopaedic treatment.[8] Records based on voxels should be used to evaluate the accuracy of orthognathic surgery due to a lower possibility of human error.[6]

In our systematic review, 10 studies used the surface as the recording method,[7,11,14,15,19,22,24,27,29,30] nine used the voxel method,[3,9,10,12,16,17,21,23,25] one used reference points,[18] and six others reported no registration method.[4,5,13,20,26,28]

In most published studies the success criteria adopted to evaluate the accuracy of VSP were defined up to 2 mm for linear differences,[3,9,12,14,16–18,20,22,27] and up to 4° for angular differences.[9,17,20,27] In two studies, the success criterion for linear differences was defined as being up to 1 mm[7,26] because VSP using a printed surgical guide was accurate and reliable. However, according to Shaheen et al,[10] errors of 2 mm or 4° are clinically insignificant. On the other hand, Tran et al[17] used a 1 mm discrepancy limit for the maxillary dental midline position, since it was considered clinically imperceptible.

In four studies, inadequate accuracy of VSP was found in posterior vertical height of the maxilla and pitch rotation.[5,7,21,31] Although Bengtsson et al[4] found a great degree of accuracy, they observed that it was poor for mandibular points. Zavattero et al[7] also reported a high degree of accuracy, but discrepancies between selected points showed a moderate degree of accuracy in two cases. The authors concluded, however, that the analysis should be interpreted with caution due to the low number of patients. Additionally, VSP contributed to a reduction in surgical time in two studies.[3,13]

Several factors hinder precise maxillary repositioning with a surgical guide, including laboratory-related preoperative errors, such as moulding, guide fabrication, and premature contacts during the MMB (maxillomandibular block). Accuracy can also be affected by the muscles of the facial skeletal, especially those of chewing, which are relaxed under general anaesthesia, the presence of TMJ disorder, and factors related to the surgeon's technical ability.[32] Shirota et al[33] reported that the accuracy of maxillary repositioning is influenced by the position of the mandibular condyle in the temporal fossa, but movement of the TMJ is unstable under general anaesthesia and may compromise the accuracy of the surgical result. Surgeons cannot guarantee that the planned osteotomy is identical to that performed in the patient, and bony remodelling can occur in the postoperative period in places close to the osteotomy line.[19]

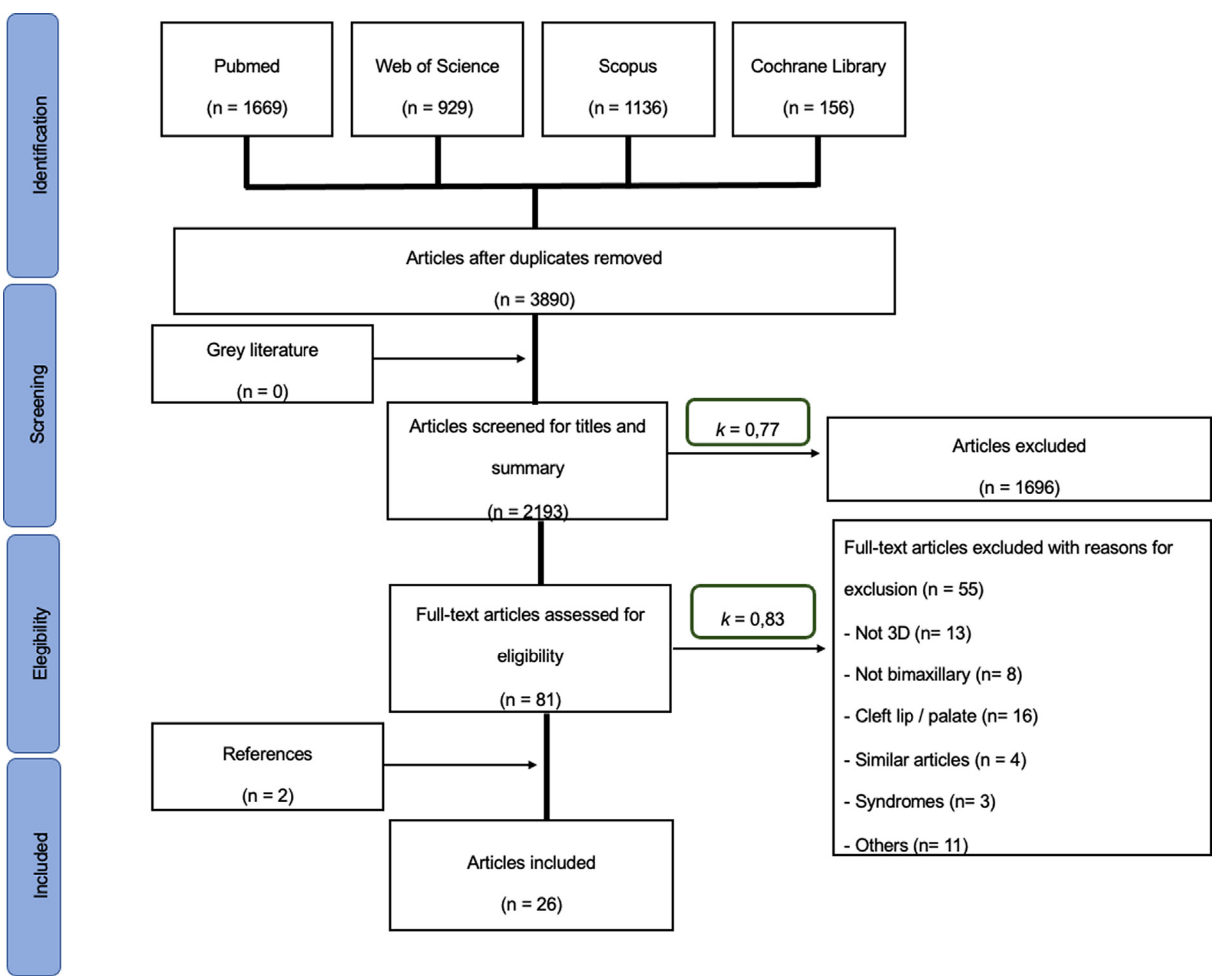

Fig. 1. Flowchart to show searches and selection criteria (adapted from PRISMA).

Table 2
Characteristics of included studies.

Author	Country	Study design	Sample	Average age (range)	Sex	Imaging	Software used in planning
Chang et al, 2020[3]	Taiwan	Prospective	30	20.6 (17.5–25.3)	F: 18, M: 12	CT	Rhinoceros (Robert McNeel & Associates)
Tonin et al, 2020[9]	Brazil	Retrospective	70	30.4 (18–50)	F: 47, M: 23	CBCT	Dolphin Imaging software version 11.95 (Dolphin Imaging & Management Solutions®)
Shaheen et al, 2019[10]	Belgium	Retrospective	15	29.6 (NI)	F: 11, M: 04	CT and CBCT	PROPLAN software (Materialise)
Kim et al, 2019[11]	Republic of Korea	Retrospective	13	22.9 (18–29)	F: 07, M: 06	CBCT	FaceGide® system
Marlière et al, 2019[12]	Brazil	Retrospective	25	27 (NI)	F: 12, M: 13	CBCT	Dolphin Imaging 11.7 Premium and Management Solutions®
Schneider et al, 2019[13]	Germany	Prospective	21 (9 3D, 12 2D)	31.1 (23–52.1)	F: 12, M: 09	CBCT	Dolphin Imaging 11.9 Premium and Management Solutions®
Wilson et al, 2019[5]	United States	Retrospective	100	21.7 (15 – 47)	F: 43, M: 57	CBCT	Mimics software (Materialise, Leuven, Belgium)
Zavattero et al, 2018[7]	Italy	Prospective	17	24.94 (18–49)	F: 9, M: 8	CT e CBCT	Proplan CMF software (Materialise CMF US, Plymouth, MI)
Ko et al, 2018[14]	Taiwan	Prospective	34	23.4 (18.1–33)	F: 19, M: 15	CBCT	Software Simplant® O&O (Materialise Dental N.V., Leuven, Belgium)
De Riu et al, 2018[15]	Italy	Retrospective	49	26.4 (NI)	F: 30, M: 19	CBCT	Maxilim®, Medicim, Nobel Biocare Group, Mechelen, Belgium
Borba et al, 2018[16]	Brazil	Retrospective	46	NI	NI	CBCT	Dolphin Imaging and Management Solutions, Chatsworth, CA, USA
Bengtsson et al, 2018[4]	Sweden	Prospective	30 (both 2D and 3D)	NI (18–30)	NI	CT	Simplant® PRO 12.02 OMS (Materialise corp., Leuven, Belgium)
Tran et al, 2018[17]	Thailand	Retrospective	15	27 (20–35)	F: 10, M: 5	CBCT	Simplant O&O (Materialise Dental NV, Leuven, Belgium)
Ritto et al, 2018[18]	Brazil	Retrospective	30 (15 2D, 15 3D)	NI	3D – F: 5, M: 10; 2D – F:8, M: 7	CT	Dolphin Imaging and Management Solutions, Chatsworth, CA, USA
Lin, 2017[19]	China	Retrospective	15	21.5 (NI)	F: 11, M: 4	CT	SimPlant Pro 11.04 (Materialise Dental, Leuven, Belgium)
Zhang et al, 2016[20]	China	Prospective	30	21.8 (19–27)	F: 14, M: 16	CT	Dolphin Imaging 11.7 Premium (Dolphin Imaging and Management Solutions, Chatsworth, CA, USA) and Mimics software (version 10.01; Materialise, Leuven, Belgium)
Baan et al, 2016[21]	Netherlands	Prospective	10	26.5 (17–45)	F: 6, M: 4	CBCT	Maxilim software (Medicim NV, Mechelen, Belgium)
Stokbro, 2016[22]	Denmark	Retrospective	30	23.1 (18–42)	F: 20, M: 10	CBCT	Dolphin 3D (Dolphin Imaging and Management, Chatsworth, CA, USA)
Zinser e Zoeller, 2015[23]	Germany	Prospective	18 (3D 08, 2D 10)	NI	NI	CT and CBCT	SimPlant, Materialise, Leuven, Belgium, and ZIB-Amira, Berlin, Germany
Badiali et al, 2015[24]	Italy	Prospective	15	24.7 (17–49)	F: 12, M: 3	CBCT	SurgiCase CMF 5.0 (Materialise, Leuven, Belgium)
Sun et al, 2013[25]	Belgium	Prospective	15	NI	NI	CBCT	Amira medical imaging software tool (Visage Imaging, Germany)
Li et al, 2013[26]	China	Prospective	6	NI (19–30)	F: 2, M: 4	CT	SurgiCase CMF 5.0 (Materialise, NV Leuven, Belgium)
Hsu et al, 2013[27]	United States	Prospective	65	26.7 (15–51)	F: 34, M: 31	CT	Simplant OMS (Materialise Dental Inc, Glen Burnie, MD)
Shehab et al, 2013[28]	Egypt	Prospective	6	23.5 (18–30)	F: 12, M: 3	MSCT	Onyx Ceph 2/6/24 (Image Instruments GmbH, Göttingen, Germany) and Voxim (IVS Solutions, Chemnitz, Germany)
Hernandez-Alfaro e Guijarro-Martine, 2013[29]	Spain	Prospective	6	23.7 (19–37)	F: 3, M: 3	CBCT	Simplant PRO OMS (Materialise Dental, Leuven, Belgium)
Zinser et al, 2012[30]	Germany	Prospective	8	22.4 (19–35)	F: 4, M: 4	CBCT e MSCT	Simplant PRO Crystal (Materialise Dental, Leuven, Belgium)

Table 3
Included studies characteristics.

Author	Registration method	Precision planning method	Success criterion	Accuracy	Main conclusion
Chang et al, 2020[3]	Voxel	Linear and angular mean differences	Linear and angular discrepancies less than 2 mm or 2 degrees, respectively	Linear – x: 2.11 mm, y: 1.38 mm, z: 0.47 mm; Angular: 1.16 °	Planning accuracy within the acceptable range
Tonin et al, 2020[9]	Voxel	Linear and angular mean differences	Mean linear difference <2 mm and mean angular difference <4 °	Linear – x: 0.26 mm, y: 0.19 mm, z: 1.31 mm; Angular – pitch: 0.54°, roll: 0.12°, yaw: 0.57°	All points had values within the range considered clinically irrelevant (<2 mm and <1 °)
Shaheen et al, 2019[10]	Voxel	Maxilla rotational and translational difference	NI	translational – x: 0.9 mm, y: 1.2 mm, z: 1.1 mm; rotational – pitch: 1.95°, roll: 0.8°, yaw: 1.6°	High accuracy. 2 mm or 4 degrees errors are clinically insignificant
Kim et al, 2019[11]	Surface	Maxilla mean linear differences	NI	ML 11/21: 0.82 mm, 13 e 23: 0.81 mm, 16: 1.19 mm, 26: 1.02 mm, ENA: 0.88 mm, ENP: 1.66 mm, APoint: 0.86 mm	Maxilla repositioning showed clinically good accuracy
Marlière et al, 2019[12]	Voxel	Difference in 3D error between VSP and surgical outcome	Deviations ≤2 mm	3D Error: 1.27 mm	This study showed an average 3D error within the clinical standards of success, less than 2 mm
Schneider et al, 2019[13]	NI	SNA, SNB and ANB mean angular differences	NI	Virtual – SNA :0.6°, SNB: 0.7°, ANB: 0.5°; Conventional: SNA: 1.5°, SNB: 1.7°, ANB: 1.6°	In the VSP there was a notable reduction in surgical time together with an accuracy Improvement
Wilson et al, 2019[5]	NI	Linear and angular mean differences	NI	Linear – x: 1.5 mm, y: 2.47 mm, z: 1.92; Angular: SNA 0.04°, SNB 0,65°, ANB 0,01°	The study showed a high degree of accuracy between the VSP and postoperative result. However, some incongruity was seen, vertically (maxilla) and sagittal (mandible)
Zavattero et al, 2018[7]	Surface	Mean linear differences	Linear differences <1 mm	11: 0,76 mm, 16: 0,54 mm, 26: 1,06 mm, Pgonion: 2,30 mm, Menton: 2,11 mm, Gonion: 0,67 mm, Ramus: 0,51 mm Condyle: 0,16 mm	A high general degree of accuracy between VSP and postoperative result was found. However, discrepancy in 2 cases showed only a moderate accuracy degree.
Ko et al, 2018[14]	Surface	Linear and angular mean differences	2 mm accuracy was considered a clinically acceptable error	Thirteen out of fifteen linear differences: <1mm, Two: <2mm; Two of three angular differences: <1 °, One: 2.04 °	Accuracy was inadequate in the posterior vertical height of the mandible and in pitch rotation
De Riu et al, 2018[15]	Surface	Linear and angular mean differences	NI	Linear: 1.98 mm, Angular: 1.19°	Virtual surgical planning showed good accuracy in 12 out of 15 evaluated parameters
Borba et al, 2018[16]	Voxel	Mean linear differences of tooth 11	Discrepancy between 0 and 2 mm	y: 0.76 mm, z: 0.96 mm	3D planning showed good accuracy
Bengtsson et al, 2018[4]	NI	Linear and angular mean differences	NI	3D – Linear: 2.78 mm; Angular: 1.06° 2D – Linear: 3.22 mm; Angular: 1.13°	Study indicated an equally high accuracy for both methods (2D/3D) studied. However, in those patients with facial asymmetry, the 3D technique had an obvious advantage
Tran et al, 2018[17]	Voxel	Linear and angular mean differences	2 mm for linear difference and 4 ° for angular difference	Linear: 0,88 mm (0,79 mm for maxillaand 1 mm for mandible); Angular: 1,16 °	Virtual planning was successfully transferred to surgery, within acceptable clinical criteria
Ritto et al, 2018[18]	Points	Mean linear differences	Mean linear difference <2 mm	3D – x: 0.90mm, y: 0.95 mm, z: 1.44 mm, 2D – x: 1.16 mm, y: 1.35 mm, z: 1.30 mm	There was no statistically significant difference between 2D and 3D planning regarding accuracy.
Lin, 2017[19]	Surface	3 points mean linear differences in maxilla: midline, 16 and 26	NI	Medium line: 1.22mm, 16: 1.23mm, 26: 1.32mm	method accuracy was acceptable in patients with normal TMJ
Zhang et al, 2016[20]	NI	Linear and angular mean differences	2 mm for linear difference and 4 ° for angular difference	Linear – maxilla: 0,71 mm, mandible: 0,91 mm; Angular: 0.95°	Virtual surgical planning provided precise repositioning in the bimaxillary orthognathic surgeries. The linear and angular difference was clinically acceptable

Baan et al, 2016[21]	Voxel	Linear and angular mean differences	NI	Linear – y: 1.41 mm, x: 0.49 mm, z: 1.85 mm; Angular- pitch: 2.72°, roll: 1.04°, yaw: 0.97°	Study showed good accuracy, however, pitch showed the greatest discrepancy between 3D planning and surgical outcome
Stokbro, 2016[22]	Surface	Average linear and rotational differences	Difference of less than 2 mm between VSP and surgical outcome	Linear – Maxilla: 0.25 mm, Mandible: 0.18 mm; Rotational- Maxilla: 0.24 mm, Mandible: 0.10 mm	Study showed a high degree of accuracy in VSP
Zinser e Zoeller, 2015[23]	Voxel	Maxilla rotational and translational difference	NI	Linear – 3D: 0.20 mm, 2D: 1.17 mm; Angular: 3D: 0.47°, 2D: 1.28°	Study showed greater accuracy in VSP when compared to conventional 2D technique
Badiali et al, 2015[24]	Surface	Mean absolute linear difference	NI	x: 1,02 mm,y: 1,19 mm, z: 0,59 mm	VSP enabled accurate maxilla repositioning post-operative results, especially in vertical dimension, which is the most challenging
Sun et al, 2013[25]	Voxel	Absolute mean difference of upper central incisor point edge	NI	x: 0.50 mm, y: 0.38 mm, z: 0.57 mm	Clinically, virtual surgical guide accuracy met the requirements for bimaxillary surgery
Li et al, 2013[26]	NI	6 maxilla reference points absolute mean difference	Maxilla position accuracy for <1mm	x: 0.7 mm, y: 0.6 mm, z: 0.8 mm	There were no significant differences between VSP and postoperative result in either direction
Hsu et al, 2013[27]	Surface	Linear and angular mean differences	Linear differences up to 2 mm and angular (rotational) up to 4 °	Maxilla – x: 0.8 mm, y: 1 mm, z: 0.6 mm, pitch: 1.5, roll: 0.9, yaw: 1.3 Mandible – x: 0.8 mm, y: 0.8	The study showed excellent accuracy for maxilla and mandible
Shehab et al, 2013[28]	NI	Mean linear differences	NI	x: 0.5 a 1.2 mm, z: 0.8 mm	The surgical guide showed great accuracy in repositioning the osteotomized maxilla to the pre-planned positions
Hernandez-Alfaro e Guijarro-Martine, 2013[29]	Surface	Mean linear differences	NI	x: 0.15 mm, y: 0.25 mm, z: 0.5 mm	This method achieved a great accuracy between VSP and surgical result.
Zinser et al, 2012[30]	Surface	Linear and angular mean differences	NI	Linear: 0.15 mm; Angular: 0.26°	Study confirmed VSP clinical and accuracy viability with surgical outcome

S (sample), CT (Computed tomography), CBCT (Cone Beam Computed tomography), MSCT (Multi Slice Computed tomography), NI (Not informed), F (Female), M (Male), ML 11/21 (Midline between the upper central incisors), 11 (upper right central incisor), 13 (upper right canine), 23 (upper left canine), 16 (upper right first molar), 26 (upper left first molar), ENA (Anterior Nasal Spine), ENP (Posterior Nasal Spine), SNA (angle between Sella, Nasion and A point), SNB (angle between Sella, Nasion and B point), ANB (angle between A point, Nasion and B point) , VSP (Virtual Surgical Planning).

Bengtsson et al[4] evaluated the accuracy of planning after 12 months. This may have been a determining factor for low accuracy in the mandible, resulting from simultaneous remodelling of the TMJ and musculature in a new maxillo-mandibular position. In addition, the sample inclusion criterion was patients with an angle class III occlusion with a minimum of 5 mm negative horizontal overjet. This may have compromised the accuracy of 3D and 2D linear measurements, which were 2.78 mm and 3.22 mm, respectively.[4]

According to Hsu et al,[27] postoperative CT scans that showed the actual surgical results were obtained within six weeks of the surgical procedure. This interval was selected to avoid bias caused by possible growth or orthodontic movement.

Zhang et al[20] found linear differences between the VSP and surgical result of 0.71 mm for the maxilla and 0.91 mm for the mandible. They also reported that virtual planning seemed to work better in the maxilla than the mandible, corroborating the findings of Tran et al,[17] who reported linear differences of 0.79 mm and 1 mm for the maxilla and mandible, respectively, and those of Hsu et al[27] However, Stokbro et al[22] showed greater accuracy in the mandible than the maxilla.

According to MINORS, the bias risk in the non-comparative studies ranged from 8 - 14 points on a scale ranging from 0-18, with an overall average of 10 points, showing a moderate risk of bias. On the other hand, the bias risk in the comparative studies ranged from 14 - 16 points on a scale ranging from 0 - 24, and presented an overall average of 15 points, also a moderate risk of bias.

Despite the attempt of Gaber et al[6] to discover a protocol for the assessment of accuracy in virtual planning for orthognathic surgery, they did not find one that was standardised. Moreover, due to the different software used in virtual planning, standardisation becomes even more difficult.

Conclusion

Virtual planning in bimaxillary orthognathic surgery (in bony tissue) that is verified by linear and angular mean differences in both the maxilla and mandible is accurate. Thus, virtual planning with a printed 3D surgical guide must be used. However, there is still a lack of consensus on standardisation, and future studies on the implementation of an evaluation protocol are still needed to confirm the overall accuracy of virtual planning.

Acknowledgements

We thank our colleagues from São Leopoldo Mandic.

Conflict of interest

We have no conflicts of interest.

Table 4
Quality assessment of included non-comparative studies according to MINOR.

	Bengtsson et al, 2018[4]	Tran et al, 2018[17]	Schneider et al, 2019[13]	Tonin et al, 2020[9]	Zinser e Zoeller, 2015[23]	Chang et al, 2020[3]	Hernandez-Alfaro e Guijarro-Martinez, 2013[29]	Hsu et al, 2013[27]
1. A clearly stated objective	2	2	2	2	2	2	2	2
2. Inclusion of consecutive patients	0	0	0	2	0	2	2	2
3. Prospective data collection	2	2	2	2	2	2	2	2
4. Outcomes suitable to study objectives	2	2	2	2	2	2	2	2
5. Study outcome impartial assessment	0	0	0	0	0	0	0	0
6. Follow-up period appropriate to the study objective	2	2	2	2	2	2	2	2
7. Follow-up loss less than 5%	2	2	2	2	2	2	2	2
8. Prospective study size calculation	0	0	0	2	0	0	0	0
Additional criteria in comparative studies case.								
9. An adequate control group	2	2	2	-	2	-	-	-
10. Contemporary groups	0	0	0	-	0	-	-	-
11. Baseline equivalence of groups	2	2	0	-	0	-	-	-
12. Adequate statistical analysis	2	2	2	-	2	-	-	-
Total Score	**16**	**16**	**14**	**14**	**14**	**12**	**12**	**12**

	Ko et al, 2018[14]	Zavattero et al, 2018 [7]	Zhang et al, 2016[20]	Baan et al, 2016[21]	Borba et al, 2018[16]	De Riu et al, 2018[15]	Kim et al, 2019[11]	Li et al, 2013[26]
1. A clearly stated objective	2	2	2	2	2	2	2	2
2. Inclusion of consecutive patients	2	2	2	2	0	2	0	0
3. Prospective data collection	2	2	2	1	0	0	2	2
4. Outcomes suitable to study objectives	2	2	2	2	2	2	2	2
5. Study outcome impartial assessment	0	0	0	0	0	0	0	0
6. Follow-up period appropriate to the study objective	2	2	2	2	2	2	2	2
7. Follow-up loss less than 5%	2	2	2	2	2	2	2	2
8. Prospective study size calculation	0	0	0	0	2	0	0	0
Additional criteria in comparative studies case								
9. An adequate control group	-	-	-	-	-	-	-	-
10. Contemporary groups	-	-	-	-	-	-	-	-
11. Baseline equivalence of groups	-	-	-	-	-	-	-	-
12. Adequate statistical analysis	-	-	-	-	-	-	-	-
Total Score	**12**	**12**	**12**	**11**	**10**	**10**	**10**	**10**

	Shehab et al, 2013[28]	Stokbro, 2016 [22]	Zinser et al, 2012[30]	Sun et al, 2013[25]	Badiali et al, 2015[24]	Lin, 2017[19]	Marlière et al, 2019[12]	Ritto et al, 2018[18]	Shaheen et al, 2019[10]	Wilson et al, 2019 [5]
1. A clearly stated objective	2	2	2	2	2	2	2	2	2	2
2. Inclusion of consecutive patients	0	2	0	0	0	0	0	0	0	0
3. Prospective data collection	2	0	2	1	0	0	0	0	0	0
4. Outcomes suitable to study objectives	2	2	2	2	2	2	2	2	2	2
5. Study outcome impartial assessment	0	0	0	0	0	0	0	0	0	0
6. Follow-up period appropriate to the study objective	2	2	2	2	2	2	2	2	2	2
7. Follow-up loss less than 5%	2	2	2	2	2	2	2	2	2	2
8. Prospective study size calculation	0	0	0	0	0	0	0	0	0	0
Additional criteria in comparative studies case										
9. An adequate control group	-	-	-	-	-	-	-	-	-	-
10. Contemporary groups	-	-	-	-	-	-	-	-	-	-
11. Baseline equivalence of groups	-	-	-	-	-	-	-	-	-	-
12. Adequate statistical analysis	-	-	-	-	-	-	-	-	-	-
Total Score	**10**	**10**	**10**	**9**	**8**	**8**	**8**	**8**	**8**	**8**

Items are scored: 0 (not reported), 1 (reported inappropriately) or 2 (reported and appropriate). The ideal global score is 16 for non-comparative studies and 24 for comparative studies.

Ethics statement/confirmation of patients permission

Since the manuscript is a systematic review, there was no need to obtain ethics approval or patients' permission/consent.

References

[See online]