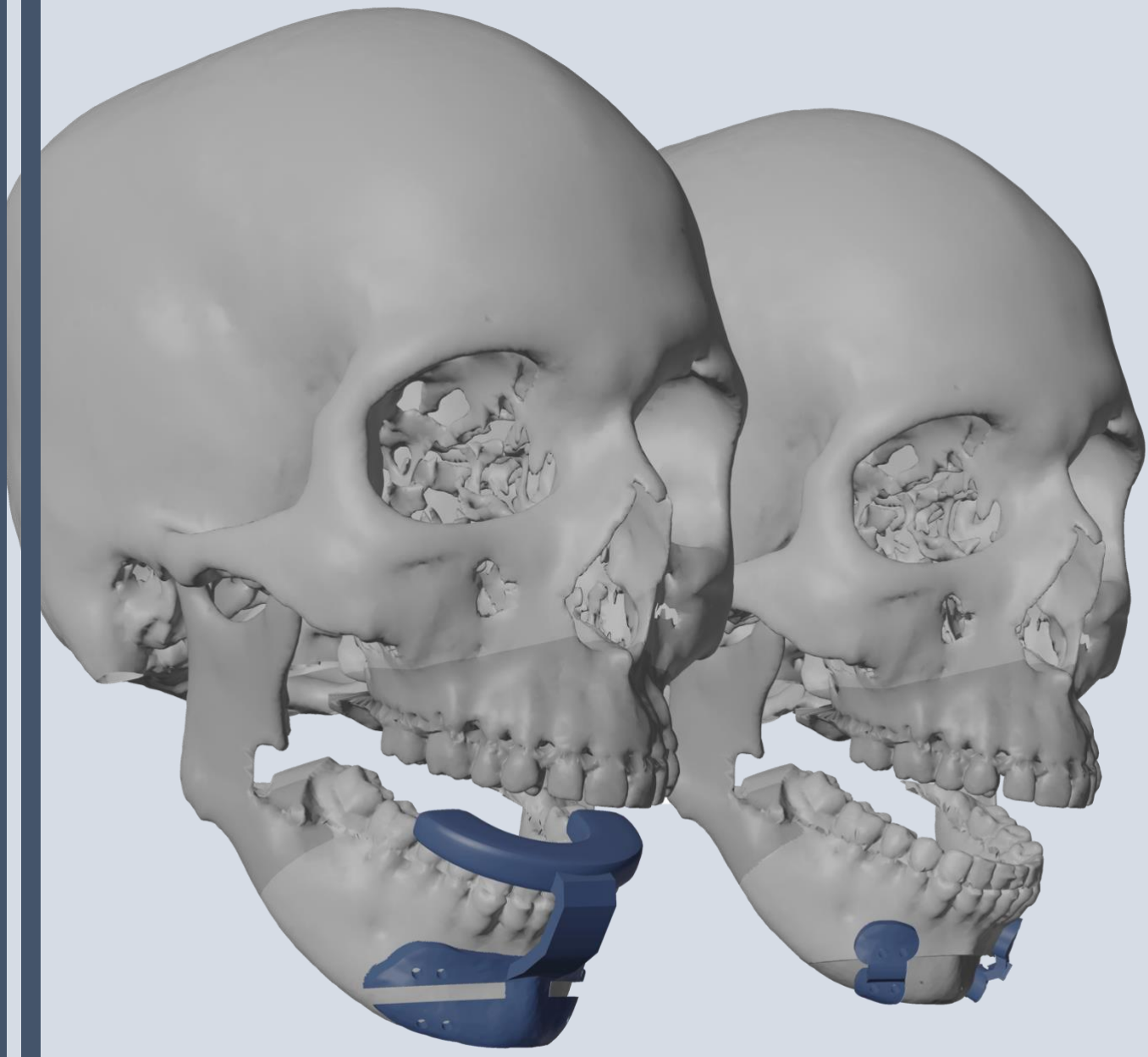
Validation of 3D-printed guidance systems for genioplasties in orthognathic surgery

A retrospective validation study &

preparation for a multicenter intervention study



J.F. Sabelis

Amsterdam UMC

TU Delft

Radboudumc

Validation of 3D-printed guidance systems for genioplasties in orthognathic surgery

A retrospective validation study & preparation for a multicenter intervention study

J.F. Sabelis
Student number: 4517377
28-08-2020

Thesis in partial fulfilment of the requirements for the joint degree of Master of Science in

Technical Medicine

University Leiden; Delft University of Technology; Erasmus University Rotterdam

Master thesis project (TM30004; 35 ECTS)
Dept. of Biomechanical Engineering, TU Delft
January 2020 – August 2020

Thesis supervisors:
Prof. Dr. J. Dankelman - TU Delft
Prof. Dr. A.G. Becking - Amsterdam UMC
Prof. Dr. T.J.J. Maal - Radboudumc
Drs. T.C.T van Riet - Amsterdam UMC
Drs. R. Schreurs - Amsterdam UMC

Thesis committee:
Prof. Dr. ir. J. Harlaar – TU Delft (chair)
Prof. Dr. J. Dankelman - TU Delft (technical supervisor)
Prof. Dr. A.G. Becking - Amsterdam UMC (medical supervisor)
Drs. R. Schreurs - Amsterdam UMC (daily supervisor)
Prof. dr. MD E. Wolvius – Erasmus MC (cum laude)
Dr. D. Dodou – TU Delft (cum laude)

Abstract

Introduction:

A chin deformity does not rarely accompany dentofacial deformities. A correction of the chin (genioplasty) is sometimes indicated in addition to orthognathic surgery. Computer-assisted surgery allows virtual planning in orthognathic surgery in three-dimensions (3D). The virtually planned movements will be transferred to the patients through 3D-printed devices, such as occlusal splints for the maxillomandibular complex. Computer-assisted surgery also enables the evaluation and quantification of the surgical result in 3D, through comparison of the virtual surgical planning and postoperative imaging. The OrthoGnathicAnalyser is a software tool that enables semi-automatic quantification of the surgical result for repositioning of the mandible, maxilla and rami in orthognathic surgery. The goal of this master thesis was to validate the application of 3D-printed guidance systems for the execution of genioplasties in the context of orthognathic surgery with a new version of OrthoGnathicAnalyser.

Validation study:

A new version of OrthoGnathicAnalyser was developed, to allow for analysis of the genioplasty. The effect of two factors (calculation of chin analysis and registration technique) on the precision and accuracy has been evaluated. This has led to two important conclusions which are implemented in the software: 1) The difference between the postoperative chin and the planned chin with respect to the realized mandible (instead of the planned mandible) should be calculated, to isolate the mandibular positioning error from the chin error. 2) Surface-based matching resulted in more accurate pitch values and was therefore implemented in the newest version.

To present and validate the newly developed software, a multicenter study was executed. A total of 25 patients were included in the study. The inter-observer and intra-observer reliability were evaluated. It was concluded that the reported results demonstrated an excellent reproducibility (ICC >0.92) of the quantification of the skeletal movements between two image sets by the OrthoGnathicAnalyser 2.0. By implementing the chin analysis in the software tool, the complete orthognathic surgery result could be quantified and compared to the virtual plan. The results of this study will be submitted to a scientific journal.

Multicenter intervention study:

To validate the 3D-printed guidance system for genioplasties, a multicenter randomized controlled intervention study was initiated. To acquire ethical approval of the local ethics committee, extensive preparations were required. Approval to start the study has been acquired within this graduation project and patient inclusion will start when COVID-19 circumstances allow continuation of clinical studies and regular patient care.

Discussion:

To validate 3D-printed guidance systems for the execution of genioplasties different projects were undertaken. A new version of the OrthoGnathicAnalyser was developed and validated. Advantageous of the OrthoGnathicAnalyser 2.0 was the implementation of the chin analysis, independence on any planning software and reduced manual input. Possible improvements were the implementation of automatic 3D landmarking and adapting the voxel-based matching algorithm to allow exclusion of the fixation material. A multicenter intervention study to evaluate the accuracy of the 3D-printed guidance system for genioplasties with the newly developed software was initiated. This study will enable a definitive conclusion about the effect of the 3D-printed guidance system on the accuracy of chin repositioning.

Preface

This is the report of my research about the validation of 3D-printed guidance systems for genioplasties in orthognathic surgery. This thesis is part of the Technical Medicine master program of the Delft University Technology, Leiden University and Erasmus University Rotterdam. The project started in January 2020 and was completed in August 2020. Due to the COVID-19 situation, the focus of the research project has shifted from an intervention study to a validation study.

The completion of this thesis signifies the end of my eight years as a student. I started in 2012 as a bachelor student of Medicine in Leiden. During my clinical internships, I learned that I needed a more technical challenge. Therefore, I switched to Technical Medicine in Delft. I was glad to have finished the personalized transition program after one year of hard work and dedication. During the master's program, my interest in 3D diagnostics and interventions was sparked. This interest was fueled during the corporate internship at Stryker and clinical internships at the departments of orthopedics, nuclear medicine, radiology and craniomaxillofacial surgery. These internships, combined with a sister who has had a very interesting 3D seminar by prof. dr. Thomas Maal, have led to the beginning of this thesis project in Amsterdam UMC, location AMC.

I would like to thank the education team in Delft, for giving me a chance to switch to a Master Program that did not exist yet. I would also like to thank (all) my supervisors for this project; prof. dr. (Jenny) Dankelman, for the remote guidance on the process and structure, prof. dr. (Eddy) Becking, for your time and feedback, prof. dr. (Thomas) Maal, for the introduction to the department and the rare, but pleasant, chats, and last, but not least, Tom and Ruud for the many feedback sessions and funny notes to brighten up my mood during the writing process at home.

I would also like to thank my family, friends and, especially, my boyfriend Nils. I imagine it was not always easy to support me during the intense transition year and during the corona period. In the end, you were there for me, each in your own way, and I appreciate it very much.

Juliana Sabelis
Amsterdam, 2020

Content

List of tables

List of figures

1. Introduction

1.1 Problem definition

1.1.1 The chin – Small but significant

Together with the bones and cartilage of the forehead, nose and zygomatic prominences, the chin plays an important role in the assessment of facial appearance, especially in the definition of the profile [1]. The harmony of the facial profile is determined in part by the size, shape, position and proportion of the chin with respect to other facial elements [2]. A chin deformity can be classified based on volumetric mass and spatial position. For example, the chin can be volumetrically small in one or multiple planes (horizontal, vertical or combination), which is called microgenia (see Figure 1B). What may appear to be a small chin, may very well be a chin that is normal sized, but positioned posteriorly to its harmonious position (also known as retrogenia, illustrated in Figure 1C). In addition, retrogenia can also be secondary to a growth deformity of the mandible. This is called pseudoretrogenia and is depicted in Figure 1D [2].

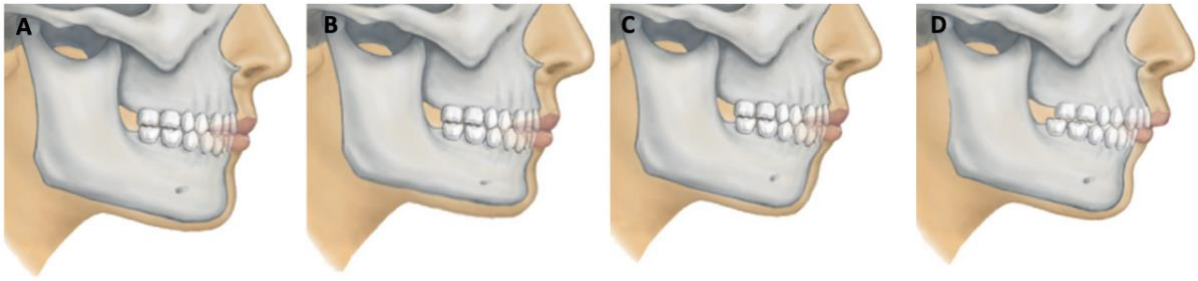


Figure 1 Visualization of the different chin classifications. A) Ideal chin for reference, B) Microgenia, C) Retrogenia, D) Pseudoretrogenia (class II mandibular retrognathia, but normal chin). From Ward et al. [2].

1.1.2 The chin – Often part of something bigger

Frequently, patients presenting with a deformity of the chin also have bimaxillary dentofacial deformities [2]. These patients may experience functional problems, such as difficulties with chewing, biting and speaking [3]. Orthognathic surgery may be performed to alleviate these functional and/or esthetic complaints. In orthognathic surgery, different techniques can be applied to adjust the skeleton to achieve functional and esthetics goals such as class I dental occlusion and facial balance and proportion [4]. Possible techniques to achieve optimal clinical results are repositioning of the maxilla and/or the mandible, by executing a Le Fort I osteotomy or bilateral sagittal split osteotomy (BSSO) respectively [4]. Depending on the patients' anatomy and pre-surgical orthodontic treatment, combinations of these techniques may be used to achieve an optimal functional and esthetic result. If necessary to achieve these goals, additional esthetic procedures such as a genioplasty (correction of the chin), rhinoplasty (correction of the nose) or zygomaplasty (correction of the cheekbones) may be used [5]. Repositioning of both jaws may produce significant changes in the patients' appearance. To guarantee a desirable outcome for every patient, an orthognathic treatment plan requires extensive consultation, psycho-social guidance and meticulous planning [6, 7]. Intraoperatively, it requires an accurate transfer method of the surgical treatment plan to the segments that need to be repositioned.

1.1.3 3D planning is key

Conventionally, surgical planning is based on two-dimensional (2D) imaging. With the availability of virtual three-dimensional (3D) imaging and software, computer-assisted surgery (CAS) has become a superior alternative to conventional planning [8]. CAS is a broad concept that encompasses different techniques. The main characteristic is that part of the surgical process is supported by computer technology [9].

In the CAS workflow, a virtual model of the patient is created. This consists of a 3D composite model of the patient which accurately represents the skeleton, dentition and the facial soft tissue [9]. The virtual model can be viewed from every angle and individual parts of the model can be masked to acquire a free view on the 3D model of interest. For example, by changing the perspective and masking the mandible and soft tissue, a free caudocranial view on the maxilla can be acquired. This view allows exact diagnosis of yaw deformities, which would not have been detected with the conventional planning methods and are very difficult to diagnose in a clinical examination [10]. This example shows the increased value of CAS in terms of diagnostics. CAS has also increased the planning potential by enabling the execution of virtual osteotomies. This allows for precise translation and rotation of the mandible, maxilla and chin in all three dimensions. Translation can be in the mediolateral (x-axis), anteroposterior (y-axis) and superior-inferior (z-axis) direction. Rotations are defined around the x-axis (pitch), around the y-axis (roll) and around the z-axis (yaw). These transformations are illustrated in Figure 2.

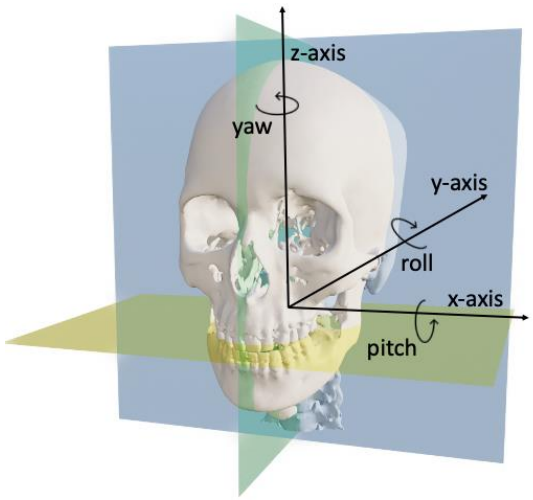


Figure 2 Possible 3D translations and rotations. These are defined as: x-axis increases from patient right towards patient left, y-axis increases from patient front towards patient back, z-axis increases from patient feet towards patient head [11] . Pitch is a rotation around the x-axis (positive is counterclockwise) and can be seen in the sagittal plane (green). Roll is a rotation around the y-axis (positive is counterclockwise) and can be seen in frontal plane (blue). Yaw is a rotation around the z-axis (positive is counterclockwise) and can be seen in the axial plane (yellow).

A virtual prediction of the outcome in soft tissue changes due to bony transformations can be made, which aids both the surgeon and patient in difficult clinical decisions. For example, the effect of the planned movements of the maxillomandibular complex on the position and orientation of the chin can be evaluated during the virtual planning. It might reveal the chin is already in midline of the face, resolving the need for a genioplasty. In other cases, the 3D planning can show the chin requires a translation and rotation. Thus, the conventional plan (‘4 mm advancement, 1 mm to the left’) is converted by the 3D planning to translational and rotational movements in three dimensions. This potential increase in complexity of the

surgical plan poses a problem for the surgeon, who has to transfer the detailed plan to the patient during the surgery.

1.1.4 3D transfer to the operation room

Distinct methods of transferring the increasingly precise surgical planning to the operation room (OR) are being developed. One method to transfer the surgical planning of the chin is a recently introduced 3D-printed guidance system [12]. The system consists of separate cutting and repositioning guides. The cutting guides are designed to guide the surgeon in performing the osteotomy (Figure 3A). As the name implies, the repositioning guides guide the surgeon in repositioning the chin segment and enable fixation with a titanium osteosynthesis plate (Figure 3B).

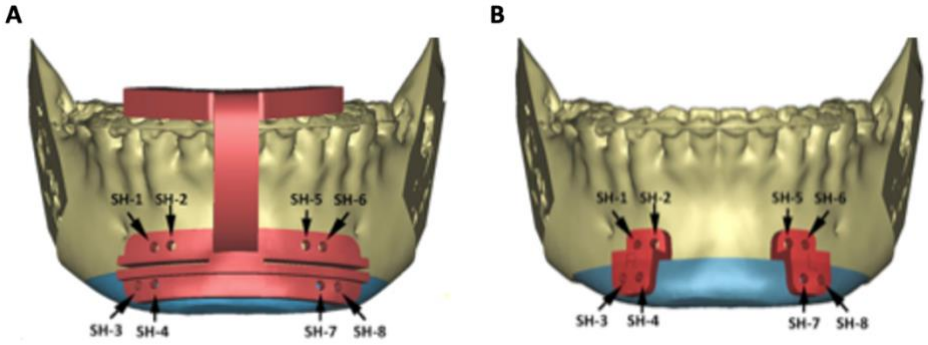


Figure 3 3D-printed guidance system of Li et al. [12]. A) Cutting guide, B) Repositioning guides, between these the titanium osteosynthesis plate can be used for fixation.

1.1.5 3D evaluation – Completing the circle

The goal of orthognathic surgery is to achieve optimal functional and cosmetic outcomes by correcting craniomaxillofacial deformities [3]. While the virtual surgical plan simulates the most ideal result, a surgeon strives to execute this plan as accurately as possible with the help of (3D-printed) tools. A small error is associated with transferring the virtual plan into clinic. Learning from the objective evaluation of postoperative results could help to improve outcomes for future patients.

When a pre- and postoperative (Cone-Beam) Computed Tomography ((CB)CT) scan are performed, evaluation of the outcomes can be performed in 3D. In 3D evaluation, the difference between the postoperative outcome and the virtual surgical plan is quantified. This can be used to compare different transfer techniques and determine which achieves the highest accuracy, precision and predictability. In addition, 3D evaluation can be used to compare the overall differences within a patient population. This enables identification and tackling of systematic errors. For example, if the chin systematically ends up off the midline of the face, 3D evaluation can tell whether this deviation is caused by achieving the planning or not. When the plan is not achieved, this could mean the design of the guide must be altered. If it is achieved, this means something else, such as soft tissue, causes the deviation. 3D evaluation enables learning from past experiences and can steepen the learning curve of the surgeon.

1.1.6 Previous work

1.1.6.1 Accuracy of genioplasties with 3D-printed guides

Only two clinical studies so far have compared the surgical outcomes of a genioplasty between 3D-printed guidance systems and conventional techniques in terms of accuracy [9, 12]. In the

study of Hsu et al., a significantly higher accuracy was reported in the experimental group (genioplasty executed with 3D-printed guides, n = 8) compared to the conventional group (genioplasty executed without 3D-printed guides, n = 16) for both the translational (maximum difference between the two groups of 0.9 mm in anteroposterior direction, $p<0.001$) and rotational (maximum difference between the two groups of 3.9 degrees in pitch, $p = 0.004$) parameters [9]. Also, the study of Li et al. (n = 88) has reported superior accuracy for the experimental group (maximum difference between the two groups of 0.58 mm in superior-inferior direction and 2.49 degrees in pitch, no p-values were mentioned) [12].

Both studies reported an inferior accuracy of the conventional methods for genioplasties, supporting the application of the 3D-printed guidance system. The study designs were however sensitive to bias, as they were not randomized. Also, the study of Hsu et al. has a relatively small sample size (included only 24 patients undergoing genioplasties) and used a different assessment than Li et al, impeding comparisons.

1.1.6.2 3D-evaluation tools

The systematic review of Haas et al. reported that, based on 3D evaluation, the current literature provides a strong indication that CAS is more accurate than classic planning [13]. Real comparisons between the different studies could not be made since the evaluation methods used to assess the accuracy were not reproducible [13]. Another review by Gaber et al. reported a lack of consensus regarding assessment and validation methods [14]. Recommendations were provided for an ideal assessment of accuracy of orthognathic surgery: 1) voxel-based registration to decrease possibility of human error, 2) automated or semiautomated evaluation of the outcome that is indicative of changes in 3D (translational or rotational based on different axes) and 3) inter-observer and intra-observer reliability should be used to validate the results.

The assessment tool that seems to meet all these recommendations is the OrthoGnathicAnalyser (OGA) as described by Baan et al. [15]. This semi-automatic 3D tool quantified the accuracy of the surgical outcomes in relation to the 3D virtual planning. The software utilizes voxel-based registration to align the preoperative (CB)CT scan to the postoperative (CB)CT scan. After limited manual input, the software calculates the differences between the surgical plan and the postoperative outcome. Outcome parameters are the difference in translation (anteroposterior, mediolateral and superior-inferior) and rotation (roll, pitch and yaw) of the mandible and maxilla. The high intraclass correlation coefficients (>0.97) indicate that the results are observer-independent [15]. Unfortunately, analysis of the chin was not implemented in this first version. Thus, objective quantification of the result of a genioplasty is not available.

1.2 Thesis structure

The goal of this research project was to validate the application of 3D-printed guidance systems for the execution of genioplasties with a new version of OrthoGnathicAnalyser. While the clinical studies, described in paragraph 1.1.6.1, suggest that the application of a 3D-printed guidance system may increase the accuracy of the execution of the genioplasty, neither supply the hard evidence necessary to prove the improvement in accuracy, as they were sensitive to bias. This lack of evidence has led to the formulation of the research question of this thesis:

“Is a prospective randomized multicenter study towards the evaluation of the accuracy of the use of 3D-printed guidance systems for genioplasties feasible?”

Part of this research project was a literature review to put 3D-printed guidance systems for genioplasties in the context of orthognathic surgery and to give an overview of the current developments of the guidance systems. Even though this review was already graded, it was chosen to include it in Appendix A, as it has been used to support decisions made in this thesis.

In order to answer the research question, a new version of the existing OrthoGnathicAnalyser needed to be developed to enable evaluation of the genioplasty itself. This would result in an assessment tool which calculates the difference in translation and rotation between the planning and postoperative outcome consistently and accurately. The new version of the software needed to be validated. For this purpose, a multicenter validation study with scan data of patients before and after bimaxillary surgery with a genioplasty was performed. Chapter 2 first describes the software development stages and concludes with a scientific article reporting on the results of the validation study.

Once the tool was validated, it could be used to assess the effect of a 3D-printed guidance system on the accuracy of the execution of the genioplasty. For this purpose, a multicenter randomized controlled intervention study was initiated. This process consisted of different stages, such as designing the study and acquiring ethical approval. The first paragraph of chapter 3 thoroughly describes the results of these stages. The last paragraph discusses learning experiences and future work.

Chapter 4 contains a general discussion on the entire research project. The previous chapters are summarized and subjects such as the limitations and future work are discussed. Finally, chapter 5 contains the conclusions drawn from this master's thesis.

2. Validation study

2.1 Introduction

In order to accurately and consistently quantify the result of orthognathic surgery, 3D evaluation software can be used [13]. As described in the introduction of this thesis, a promising tool is the OrthoGnathicAnalyser (OGA), developed by Baan et al. [15]. This 3D tool can objectively quantify the positional changes of the mandible and maxilla due to the surgery and compare these to the virtual planning. However, analysis of the chin itself was not implemented in the previously validated version of the OGA. Since the research question of this thesis requires an accuracy assessment of the genioplasty, a dedicated chin analysis was developed and incorporated in the existing OGA workflow.

Several stages of software development were completed. These stages are explained in the following paragraph. After the software development, ethical approval was gained (see Appendix B) to start a multicenter study to validate the developed software. The results are described in a scientific article in paragraph 2.3.

2.2 Software development of chin analysis

2.2.1 Introduction

The new version of the quantification tool must be tested to verify that the results are accurate and precise. When a test has a high accuracy, it means the results are close to the ground truth. When a test has a good precision, it means the results of repeated measurements are (almost) equal [16]. Ideally, a test has a good accuracy and precision. This should hold for both repeated measurements performed by one observer and by different observers. Figure 4 illustrates the difference between accuracy and precision. Based on clinical experience, a difference between the postoperative result and planning of 1 mm and 2 degrees were defined to be clinically acceptable. These values will be used as a benchmark for the subsequent tests.

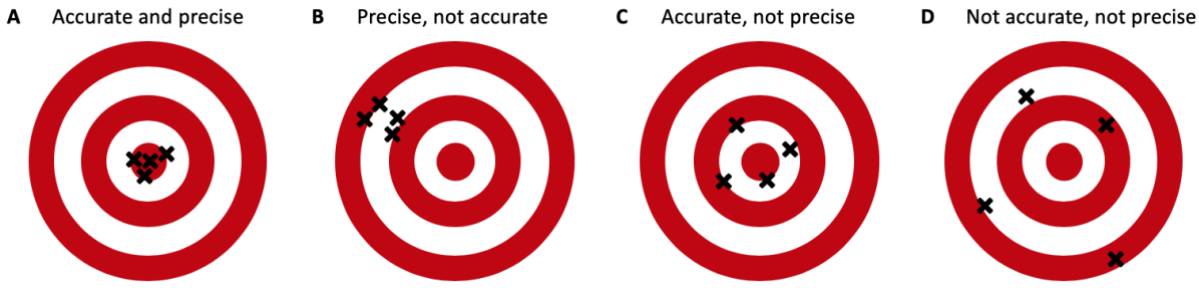


Figure 4 Difference between accuracy and precision. Bull's eye of the target represents the ground truth. The black crosses signify individual test results. A) Accurate and precise, B) Precise but not accurate, C) Accurate but not precise, D) not accurate and not precise.

To allow for analysis of the chin translation and rotation, a methodology similar to the one currently used for the mandibular and maxillary analysis, was implemented for the chin in OGA version 1.1. The primary target of the software development was to allow for an individual analysis of the chin. Secondary, the effect of two distinct factors within the workflow on the precision and accuracy were evaluated. The first factor was whether the chin analysis is performed with respect to the planned or to the realized mandible. The second factor was the registration technique applied for the chin matching.

2.2.2 Materials and methods

For the development stage, three different cases of an existing anonymous database were selected. These are described in Table 1.

Table 1 Different patient cases used for software development

	Type of surgery	Notes
Patient 1	Bimax with genioplasty	Small chin Dentition not in occlusion
Patient 2	Bimax with genioplasty	Large chin
Patient 3	Bimax without genioplasty	Ground truth

The first two patients have had bimaxillary surgery in combination with a genioplasty. These cases can be used to evaluate the precision of the software by analyzing each case multiple times and comparing the results. Since the actual movement of the chin due to surgery is not known exactly, the accuracy cannot be checked quantitatively. A visual check of the alignment of the matched preoperative models to the postoperative models could however help in evaluating the accuracy. The third patient underwent bimaxillary surgery without a genioplasty. Thus, the translation of the chin should be equal to the translation of the mandible. Voxel-based matching (VBM) has been previously validated for the registration of the mandible [17], so the resulting transformation served as the ground truth. This provided the opportunity to quantitatively assess the accuracy of the software.

2.2.2.1 Analysis with respect to planned or realized mandible

2.2.2.1.1 Testing version 1.1 of OrthoGnathicAnalyser

In version 1.1 of OGA, the user needed to indicate a landmark which was used as the origin for the evaluation. After the region of interest was set to only contain the specific segment (in this case the chin), voxel-based matching (VBM) was performed. The chin analysis was performed by calculating the difference between the postoperative chin segment and the original planned chin segment with respect to the planned mandible. To test the accuracy of this calculation, patient 1 was analyzed with the software by one observer.

2.2.2.2 Application of registration techniques

2.2.2.2.1 Comparison of registration techniques

An alternative registration technique, surface-based matching (SBM) will be compared to the previously mentioned VBM. Both techniques are explained in more detail in the paragraph “Registration Techniques” in Appendix A. To verify the accuracy and precision of the rotational outcomes of the two registration techniques for the chin, additional analyses were performed outside the OGA module, in the main 3D analysis software called 3dMedX (3D lab - Radboudumc, Nijmegen, the Netherlands). The image data of patient 3 was used to enable actual evaluation of the accuracy. One observer performed the matching of the preoperative and postoperative mandible first to establish a ground truth. Subsequently, the registration of the chin segment was performed three times with SBM and three times with VBM. Since the chin segment was not repositioned in this patient, the chin registration should equal the outcome of the mandible registration. The accuracy was evaluated by comparing the mean of the absolute difference between the repeated measurements and the ground truth. The precision was evaluated by calculating the maximum mutual difference between the individual measurements. It was impossible to reset the origin of the calculation to a landmark indicated on the chin, hence only the rotations were assessed.

2.2.2.2.2 Testing version 1.2 of OrthoGnathicAnalyser

In version 1.2 of OGA, the chin analysis was implemented with a calculation of the chin segment with respect to the realized mandible. To perform a final accuracy check, analysis of patients who had an actual genioplasty (patients 1 and 2) were analyzed with this version. Since no ground truth was available for these patients, it was decided to perform a visual check to evaluate whether the registered postoperative chin models were aligned with the postoperative hard tissue models. If the models were aligned, the results would be assumed to be accurate. These visual checks were performed by one observer.

2.2.2.2.3 Testing version 1.3 of OrthoGnathicAnalyser

In version 1.3 of OGA, SBM instead of VBM of the chin was implemented. In this stage of software development, the software was tested by one observer for the same patients as in paragraph 3.2.2, patient 1 and 2. The accuracy was verified by checking the alignment results of SBM.

2.2.3 Results

2.2.3.1 Analysis with respect to planned or realized mandible

2.2.3.1.1 Testing version 1.1 of OrthoGnathicAnalyser

In the first version of OGA, the chin analysis was based on the difference between the postoperative chin and the planned chin relative to the planned mandible. Table 3 shows the results of OGA version 1.1 for the mandible and chin.

Table 2 Results of the analysis with OrthoGnathicAnalyser version 1.1 of patient 1

		Mandible			Chin		
		Planned	*Realized*	*Difference*	*Planned*	*Realized*	*Difference*
Translation (mm)	RL	-0.3	0.2	0.5	0.0	-0.5	-0.5
	AP	-12.5	5.2	17.6	3.0	20.5	17.5
	SI	3.3	-0.9	-4.2	0.0	-0.9	-0.9
Rotation (degrees)	Roll	-0.2	-0.4	-0.2	0.0	0.3	0.3
	Pitch	-4.0	-2.7	1.3	0.0	1.5	1.5
	Yaw	0.0	-0.7	-0.7	0.0	0.4	0.4

R = right, L = left, A = anterior, P = posterior, S = superior, I = inferior

With regard to the translational results, large differences between the planned and realized values were reported for both the mandible (17.6 mm in anteroposterior direction) and the chin (17.5 mm in anteroposterior direction).

2.2.3.2 Application of registration techniques

2.2.3.2.1 Comparison of registration techniques

The rotational results of the VBM of the chin and the ground truth are reported in Table 3.

Table 3 Rotational results of voxel-based matching of the preoperative chin to the postoperative chin of an orthognathic patient without a genioplasty.

	Roll		Pitch		Yaw	
	Result	*Difference**	*Result*	*Difference**	*Result*	*Difference**
Ground truth ** (degrees)	0.06		0.05		0.54	
Mean *** (degrees)	0.15	0.16	1.22	1.17	0.88	0.34

#1	-0.04	0.10	0.14	0.09	1.31	0.77
#2	0.19	0.13	1.75	1.70	0.74	0.20
#3	0.31	0.25	1.78	1.73	0.58	0.04

* Difference is calculated by taking the absolute value of the ground truth minus the result.
** The 'ground truth' row is the result of voxel-based matching of the entire mandible.
*** The 'mean' row contains the mean values of the repeated VBM of the chin segment only.

The accuracy outcome parameter is reported in the intercept of the difference columns and the mean row in Table 4. For VBM an accuracy of 0.16, 1.17 and 0.34 degrees was computed, for the roll, pitch and yaw, respectively.

The precision was defined as the maximum mutual difference between the individual repetitions reported in the difference columns. When comparing these individual difference values, a maximum difference of 0.15 degrees roll, 1.64 degrees pitch and 0.73 degrees yaw was reported.

The rotational results of the SBM of the chin and the ground truth are reported in Table 4.

Table 4 Rotational results of surface-based matching of the preoperative chin to the postoperative chin of an orthognathic patient without a genioplasty.

	Roll		Pitch		Yaw	
	Result	*Difference**	*Result*	*Difference**	*Result*	*Difference**
Ground truth ** (degrees)	0.06		0.05		0.54	
Mean ***(degrees)	0.14	0.08	0.16	0.13	0.79	0.25
#1	0.11	0.05	0.03	0.02	0.76	0.22
#2	0.14	0.08	0.21	0.16	0.93	0.39
#3	0.17	0.11	0.25	0.20	0.67	0.13

* Difference is calculated by taking the absolute value of the ground truth minus the result
** The 'ground truth' row is the result of voxel-based matching of the entire mandible.
*** The 'mean' row contains the mean values of the repeated SBM of the chin segment only.

The accuracy of SBM was 0.08 degrees for the roll, 0.13 degrees for the pitch and 0.25 degrees for the yaw. The precision of SBM is 0.06, 0.18 and 0.26 degrees for the roll, pitch and yaw respectively.

2.2.3.2.2 Testing version 1.2 of OrthoGnathicAnalyser

The results of patient 1 are presented in Figure 5A-C. The registered preoperative chin model (in red) is shown in a coronal, sagittal and transverse view. When comparing it to the postoperative hard tissue model in Figure 5B and 5C, the yaw and roll seem to be accurate. However, in Figure 5A it can be seen that the pitch of the transformed chin model deviates.

The results of patient 2 are shown in Figure 5D-F. Again, the yaw and roll are quite accurate. Figure 5D illustrates the same issue with the pitch as seen in patient 1. In this case, the planned pitch was 0 degrees and according to the quantitative analysis a pitch of 0.3 degrees was executed. However, the gap between the chin segment and the mandible (encircled in white) suggests the realized pitch is larger.

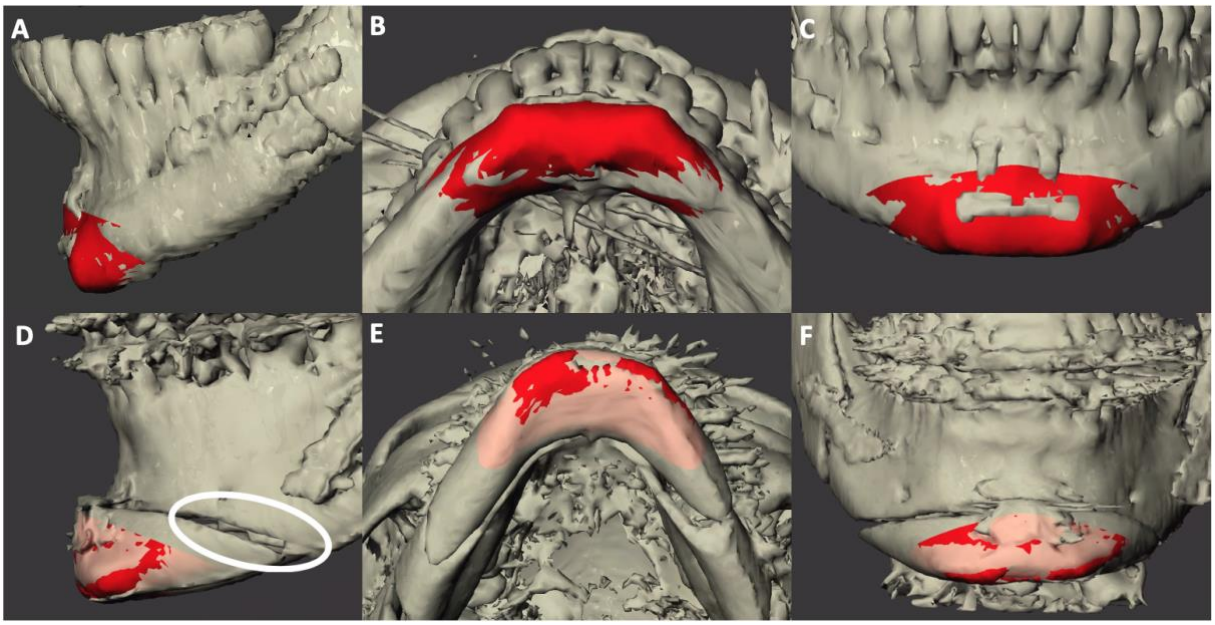


Figure 5 Visualization of alignment of the voxel-based matched models; the preoperative chin models (red) and the postoperative hard tissue model (beige). A, B and C show the matching of the preoperative chin model on the postoperative hard tissue model for patient 1. In A, it can be seen that the chin segment would require an extra clockwise rotation to align with the hard tissue model. D, E and F show the matching of the preoperative chin model (red) on the postoperative hard tissue model for patient 2. In D, the gap (encircled in white) indicated that a pitch of more than 0.3 degrees is executed.

2.2.3.2.3 Testing version 1.3 of OrthoGnathicAnalyser

The resulting models are well aligned with the postoperative hard tissue model in all three views for both patients. Figure 6A and 6D illustrate the alignment of the pitch. Quantitative analysis by OGA reported a pitch value of 6.4 degrees for patient 2, which matches the gap between the chin and the mandible.

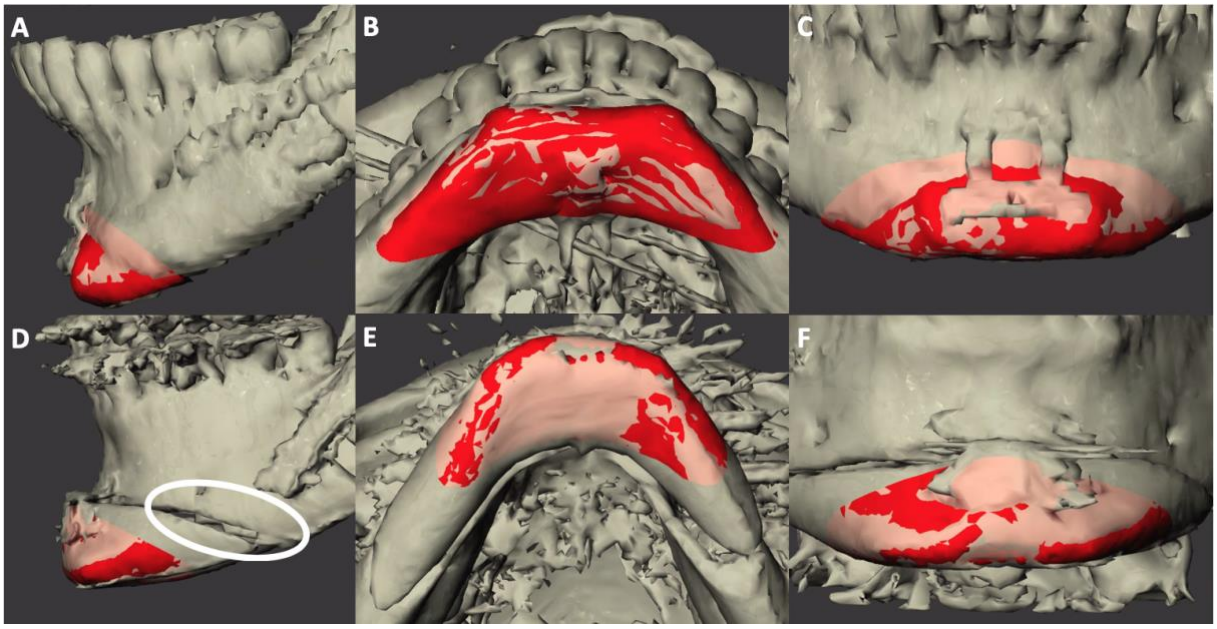


Figure 6 Visualization of alignment of the surface-based matched models; the preoperative chin models (red) and the postoperative hard tissue model (beige). A, B and C show the matching of the preoperative chin model on the postoperative hard tissue model for patient 1. In A, it can be seen that the chin segment is well aligned with the postoperative hard tissue model. D, E and F show the matching of the preoperative chin model (red) on the postoperative hard tissue model for patient 2. In D, the gap (encircled in white) matched the calculated pitch of 6.4 degrees.

2.2.4 Discussion

Throughout the entire process, considerations and conclusion were drawn based on the results found at that moment. This reasoning will be described for each individual step in the process to guide the reader through the decisions made.

2.2.4.1 Analysis with respect to planned or realized mandible

2.2.4.1.1 Testing version 1.1 of OrthoGnathicAnalyser

In the first version of OGA, the chin analysis was based on the difference between the postoperative chin and the planned chin relative to the planned mandible. The consequence of this, is that any positioning error of the mandible affected the results of the chin. Since the postoperative scan of patient 1 was performed with a suboptimal occlusion, the mandibular position differed from the planned position, especially in the translations. While it cannot be guaranteed that the error of the chin translation is zero, it was reasoned that a realized posterior translation of 20.5 mm was an overestimation of the realized translation.

In theory, the identified difference could be caused by inaccuracy in executing the surgical plan (either chin or mandible) or by inaccuracy in the condylar positioning during the postoperative scan (for example, if the patient's dentition is not in occlusion). For patient 1, the most obvious explanation was the latter. To isolate the chin translation and rotation from the mandible, it would be superior to calculate the difference between the postoperative chin and the planned chin with respect to the position of the realized mandible instead of the planned mandible. Figure 7 illustrates the difference in these calculations. In this example, the postoperative mandible does not match the preoperative planning, which is also reflected in the analysis of the chin in Figure 7D and 7E. When calculating the difference between the postoperative chin and the planned chin with respect to the realized mandible (Figure 7F and 7G), it is obvious that the chin repositioning was accurate. In the OGA version 1.1, analysis is performed in the ABCDE order, while the preference is ABCFG. It was decided to change this order to ABCFG in the OGA version 1.2.

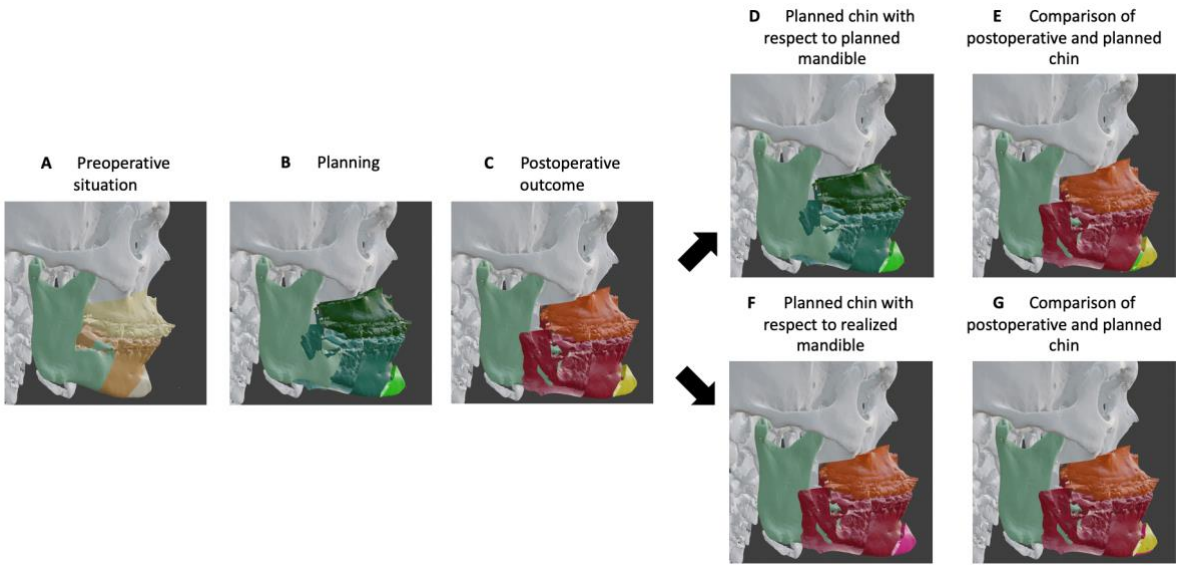


Figure 7 Illustration of the difference in calculation of the chin. A) Preoperative situation, B) Preoperative planning with the planned chin (bright green). C) Postoperative situation is simulated with realized chin (yellow). Evaluation of the result of this surgery can be performed by comparing the postoperative chin with the planned chin with respect to the planned mandible (D-E), or to the realized mandible (F-G). D) Planned chin with respect to the planned mandible. E) Compares the difference between the postoperative chin and the planned chin with respect the planned mandible. In this case, it seems the planning of the chin is not reached. F) Planned chin transformed to match the planning of the chin with respect to the

realized mandible (pink). G) Compares the difference between the postoperative chin and the planned chin with respect to the realized mandible. In this case, it is visible the planning of the chin is executed quite well.

2.2.4.2 Application of registration techniques

2.2.4.2.1 Comparison of the registration techniques

Based on the results presented in paragraph 2.2.3.2.1, the accuracy and precision of SBM seemed slightly better than that of VBM. Both techniques had acceptable accuracy and precision values below 1.0 degrees for the roll and yaw. For the pitch, VBM resulted in an accuracy and precision of 1.17 and 1.64 degrees, compared to an accuracy and precision of 0.13 and 0.18 degrees with SBM.

These values were within the predefined, clinically acceptable 2 degrees. According to literature, VBM is the registration method of preference, due to its user independency [17]. Based on these two arguments and the fact that a singular case provided insufficient evidence to alter the workflow, it was initially chosen to hold on to VBM as registration algorithm in version 1.2 of OGA.

2.2.4.2.2 Testing version 1.2 of OrthoGnathicAnalyser

The observations from the visual check led to the hypothesis that the high density of the osteosynthesis material with respect to the small volume of the chin segment affected the mutual information metric (see paragraph “Registration Techniques” in Appendix A) in VBM too much. Due to these observations it was concluded that the pitch accuracy of VBM is insufficient in case of the chin.

The data in paragraph 2.2.3.2.1 implies a sufficient accuracy and precision of the chin matching with the alternative registration method, SBM. Because SBM requires the user to manually indicate the surface of interest which will be used for the matching, the osteosynthesis material could also be excluded. Despite the conclusion in the previous paragraph (2.2.4.2.1), it was chosen to implement SBM instead of VBM for the registration of the chin.

2.2.4.2.3 Testing version 1.3 of OrthoGnathicAnalyser

By comparing Figure 5 to Figure 6, it was concluded the accuracy of SBM is better than that of VBM for the chin matching. Considering the improved accuracy, the software was deemed ready to be released and applied in the validation study (2.3).

2.3 Scientific article: “Validation of the OrthoGnathicAnalyser 2.0 – 3D accuracy assessment tool for bimaxillary surgery and genioplasty”

2.3.1 Abstract

Bimaxillary surgery is a widely performed procedure to correct dentofacial deformities. Virtual treatment planning is an important preparation step. One advantage of the use of virtual treatment planning is the possibility to assess the accuracy of bimaxillary surgery. In this study, a tool (OrthoGnathicAnalyser 2.0), which allows for quantification of the accuracy of bimaxillary surgery, is presented and validated. 30 patients who underwent bimaxillary surgery in combination with a genioplasty were selected from three different centers in the Netherlands. A pre-operative (CB)CT scan, virtual treatment planning and postoperative (CB)CT scan were required for assessing the accuracy of bimaxillary surgery. The

preoperative and postoperative (CB)CT scans were aligned using voxel-based matching. Furthermore, voxel-based matching was used to align the pre-operative maxilla, mandible and rami towards their postoperative position whereas surface-based matching was used for aligning the pre-operative chin towards the postoperative position. The alignment resulted in a transformation matrix which contained the achieved translations and rotations. The achieved translations and rotations can be compared to planning values of the virtual treatment plan. To study the reproducibility, two independent observers processed all 30 patients to assess the inter-observer variability. One observer processed the patients twice to assess the intra-observer variability. Both the intra- and inter-observer variability showed high ICC values (> 0.92) and low measurement variations (< 0.673±0.684mm and < 0.654±0.824º). The results of this study show that the OrthoGnathicAnalyser 2.0 has an excellent reproducibility for quantification of skeletal movements between two (CB)CT scans.

2.3.2 Introduction

In orthognathic surgery, suboptimal facial appearance and function may be improved by correcting dentofacial deformities [18]. Three-dimensional (3D) imaging has enhanced the potential and accuracy of the orthognathic surgery workflow [9]. The introduction of cone-beam computed tomography (CBCT) in combination with virtual imaging software enables diagnostics, planning and evaluation in 3D. This has improved quantification of, formerly difficult to measure, characteristics of dentofacial deformities. These include rotations in the axial plane (yaw) or frontal plane (roll or occlusal cant) [10]. Additive manufactured occlusal splints are based on the virtual surgical planning (VSP) and used intra-operatively to accurately execute the VSP during surgery [19]. These new 3D techniques have led to more predictable postoperative outcomes and a reduction of surgical error [20].

Similar to VSP, postoperative accuracy of the performed surgery should be evaluated in 3D. The result of orthognathic surgery was traditionally assessed in two dimensions, using pre- and postoperative lateral radiographs [17]. Contemporary software enables automatic matching, also called registration of two 3D imaging datasets. Voxel-based matching (VBM) is the registration method of preference, due to its higher accuracy and user independency [14]. In this technique, the two volumes of interest (VOI) are aligned by maximizing the overlap of the greyscale values of the individual voxels [21]. After aligning the VOI, the translations and rotations in the sagittal, coronal and axial planes (six degrees of freedom) realized by the orthognathic surgery, can be computed and compared to the VSP [8, 22]. The systematic review of Gaber et al. [14], has reviewed several 3D postoperative assessment methods of virtually planned orthognathic surgery. The OrthoGnathicAnalyser (OGA), as described in our previous study [15], was identified as the 3D assessment tool of choice, due to the application of VBM and the semi-automatic approach. Over time, the OGA has already been applied in large clinical studies [23, 24], demonstrating its clinical applicability.

After validation of the first version, the development of the OGA continued and has resulted in OGA 2.0. While the former version only enabled analysis of the mandible, maxilla, and the ramus, the new version also allows analysis of the chin segment. In addition, the efficiency of the workflow has been improved, requiring less manual interaction and computing time. The software is compatible with various VSP software, such as IPS CaseDesigner (KLS Martin Group, Tuttlingen, Germany) and Dolphin 3D (Dolphin Imaging & Management Solutions, Chatsworth, USA). The purpose of this study was to present and validate the new version of the OGA (2.0) in patients who underwent bimaxillary surgery in combination with a genioplasty. A multicenter approach was chosen to assess the robustness of the software

tool, as different centers use different imaging protocols and hardware from different manufacturers to obtain preoperative and postoperative imaging.

2.3.3 Materials & Methods

2.3.3.1 Workflow of OrthoGnathicAnalyser 2.0

The workflow of OGA 2.0 was based on the workflow described in the previous article [15] and is illustrated in Figure 8. In preparation for the surgery, the acquisition of a preoperative (CB)CT scan is required. This scan was used for the virtual planning of the subject with planning software. After the surgery, a postoperative (CB)CT scan was acquired.

The assessment of discrepancy between VSP and postoperative outcome was performed semi-automatically, using the OGA module which was implemented in the in-house developed 3D analysis software called 3dMedX (version 1.2.4.1, 3D lab Radboudumc, Nijmegen, the Netherlands). 3DMedX is a standalone software tool based on the C++ OpenInventor Toolkit (version 9.9.1.4, Thermo Fisher Scientific, Waltham, Massachusetts, USA). To start the analysis with OGA, the raw preoperative and postoperative (CB)CT scans (in Digital Imaging and Communications in Medicine (DICOM) format) were imported in the software. From the VSP the following files were necessary: the original and planned 3D models (as Standard Tessellation Language (STL) files) and a transformation matrix (in comma separated value (csv) format). The transformation matrix described the transformation of the virtual models to the natural head position (NHP) on which the VSP was based. When no transformation matrix was available, the software provided a wizard-tool to identify the NHP.

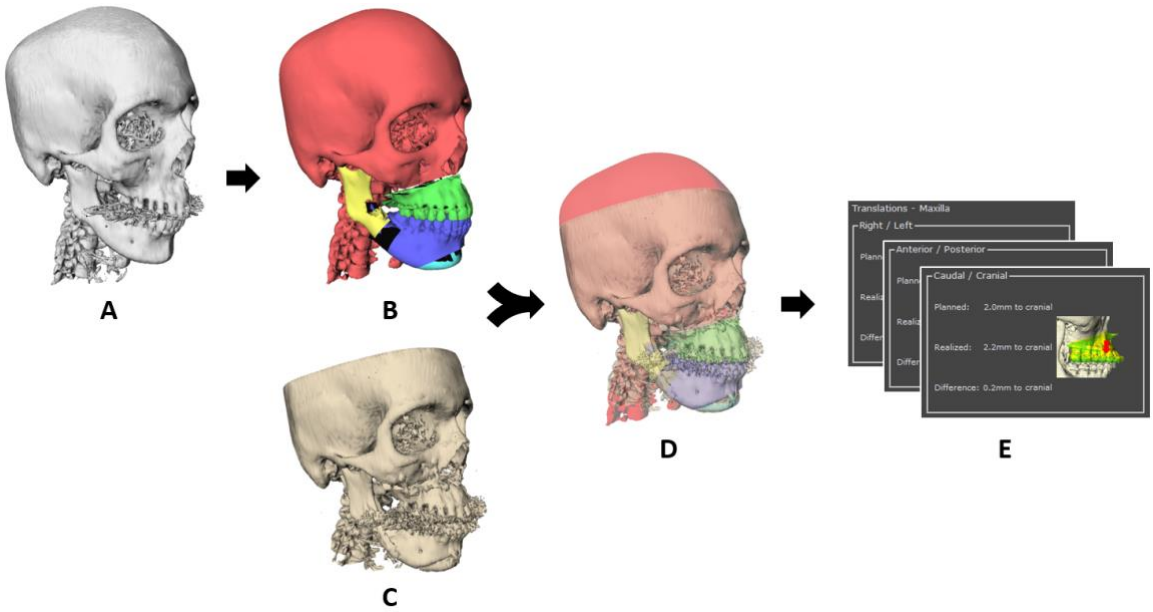


Figure 8 Global overview of the workflow of OrthoGnathicAnalyser 2.0. A) Preoperative (CB)CT scan of the patient. B) Virtually planned 3D models. C) Postoperative (CB)CT scan. The postoperative (CB)CT scan is voxel-based matched to the preoperative (CB)CT scan. Individual segments of the preoperative (CB)CT scan are matched to the postoperative (CB)CT scan. D) Overlap of the postoperative (CB)CT scan and planned STL models. E) planned, realized and difference of the three translation and three rotation parameters are computed for the maxilla, mandible, chin, left and right ramus.

Next, the user was asked to indicate four rotation points, which were used as reference points for the calculation of translations and rotations in subsequent analyses. The first point was the upper incisor point, defined as the most mesial point on the incisal edge of element 11. The second point was pogonion, as described by Swennen et al [25]. The third and fourth

points were the left and right rotation centers of the condylar head. The upper incisor point was utilized as the origin (and thus rotation point) to align the 3D models to NHP.

To compute the six degrees of freedom in VSP, the preoperative STL models were automatically matched to the planned STL models using surface-based matching (SBM). The resulting transformation matrix was calculated to represent the planned rotations and translations around the four previously indicated rotation points.

In the next steps, voxel-based matching (VBM) was used to register the individual bony segments. In VBM, a region of interest (ROI) in both scans is selected, which will be subsequently aligned based on the greyscale values [26]. First, the pre- and postoperative (CB)CT scans are aligned based on the ROI, unaffected by surgery, which consisted of the anterior cranial base, zygomatic arches and forehead [27]. For the maxilla, mandible and the left and right ramus, ROI boxes were selected to match the osteotomized bone segments. The transformation matrices, describing the translations and rotations from the preoperative models to the postoperative models, were recorded.

For registration of the chin segment, SBM was implemented instead of VBM (see Figure 9). Surface models representing the chin were generated from the preoperative and postoperative DICOM datasets. The preoperative and postoperative chin segments are roughly aligned manually, after which SBM was performed by using the unaltered caudal part of the chin, excluding the area of osteosynthesis plate. The transformation matrix obtained after SBM of the preoperative model on the postoperative model was recorded.

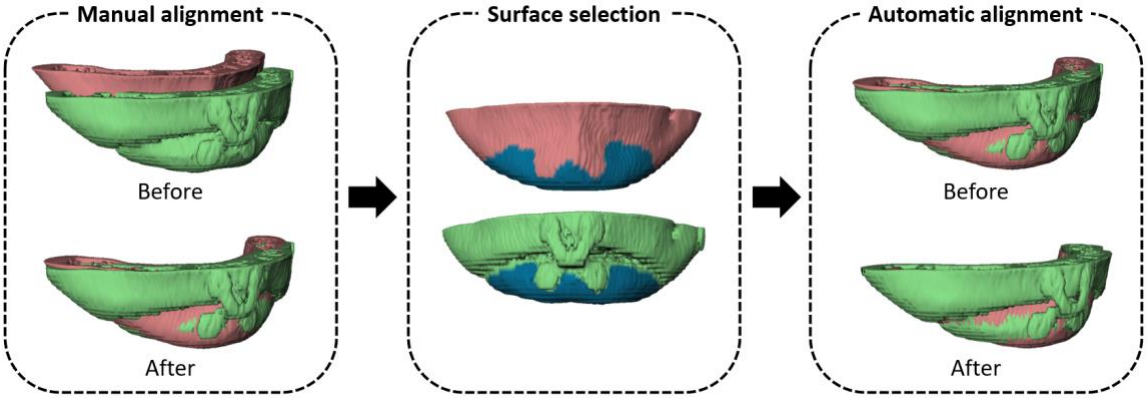


Figure 9 Surface-based matching of the chin. The preoperative chin model (red) was first manually aligned to the postoperative chin model (green). The user needed to select the surface on both models (blue). The selected surface will be used for the automatic surface-based matching.

The resulting transformation matrices were calculated to represent the surgically achieved rotations and translations around the four previously indicated rotation points for each segment. Finally, the differences between the planned and achieved movement of each segment in six degrees of freedom (translation and rotation in sagittal, coronal and axial plane) were calculated. For the chin evaluation, the difference between the planned movement of the chin and its postoperative position was calculated and corrected for the postoperative mandibular position. This excluded the potential mandibular error from the accuracy result of the chin. For the rami, only rotations were computed, assuming the condylar heads were not translated.

2.3.3.2 Validation study

Thirty subjects were enrolled in this multicenter retrospective validation study, in three centers: Amsterdam University Medical Center (UMC), Location AMC, MKA Kennemer & Meer, location Haarlem and Radboud University Nijmegen Medical Center. Per center, ten subjects with dentofacial deformities who underwent bimaxillary surgery in combination with a genioplasty between 2016 and 2020 were considered for inclusion. Availability of pre- and postoperative (CB)CT data was required. Exclusion criteria were the use of different imaging modalities (i.e. a preoperative CT scan with postoperative CBCT scan or vice versa), previous history of surgery in the maxillomandibular region and high complex cases such as multi-piece Le Fort I or cleft lip and palate cases. Prior to data analysis, all subject data were completely anonymized. The study was approved by the local ethics committee of Amsterdam UMC, location AMC (W20_127).

2.3.3.2.1 Image acquisition

The preoperative scan and postoperative scan were acquired according to the clinical protocol of the individual centers. The dental arches were recorded by making a CBCT scan of dental casts. Image acquisition parameters are described in Table 5.

Table 5 Scanning details per center

	Radboudumc	Amsterdam UMC		MKA Kennemer & Meer
	CBCT	*CT*	*CBCT*	*CBCT*
System	Imaging Sciences International I-CAT 17-19	Siemens SOMATOM Force	Planmeca ProMax	Vatech PCT-90LH
Protocol	Extended Height	Face	Skull	Dental
Potential (kV)	120	120	96	106-108
mA	5	360	10	6
FOV	17 x 23 cm	24 x 24 cm	23 x 25 cm	21 x 21 cm
Scanning time	1x 17.8 sec	1x 1 sec	1x 24 sec	1 x 18.0 sec
Voxel size	0.30 mm x 0.30 mm x 0.30 mm	0.47 mm x 0.47 mm x 1 mm	0.40 mm x 0.40 mm x 0.40 mm	0.40 mm x 0.40 mm x 0.40 mm

CBCT = cone beam computed tomography, CT = computed tomography, FOV = field of view, kV = kilovoltage, mA = milliampere

2.3.3.2.2 Surgery planning

All cases were virtually planned in IPS CaseDesigner, version 2.0.4.2 (KLS Martin Group, Tuttlingen, Germany). A 3D virtual hard-tissue and soft-tissue model were rendered and oriented in the NHP of the subject. The maxilla, mandible, chin and rami were repositioned towards their desired position. The required 3D models and transformation matrix were exported.

2.3.3.2.3 Clinical validation and evaluation

Two independent observers analyzed the (CB)CT data sets of all cases in order to validate OGA 2.0. To determine inter-observer variability, both observers performed the OGA workflow for each subject independently. For intra-observer variability, one of the observers repeated the workflow on all cases a second time and in a random order, with an interval of at least two weeks between both assessments.

2.3.3.2.4 Statistical analysis

The absolute inter-observer and intra-observer difference was calculated. One-way multivariate analysis of variance (MANOVA) was used to determine statistical differences between centers. For the evaluation of the inter-observer and intra-observer correlation and agreement, the intra-class correlation coefficient (ICC) was calculated with two-way random and two-way mixed models respectively. Statistical data analyses were performed with IBM SPSS software, version 26.0 (IBM Corp., Armonk, NY, USA).

2.3.4 Results

A total of thirty subjects (ten subjects per participating center) were included in this study. During data analysis, four subjects of the Amsterdam UMC and one subject of the MKA Kennemer & Meer were excluded because of motion artefacts (n=2), corrupt DICOM data (n=2) or incorrect field of view (n=1). This resulted in a study population of 25 subjects. The demographics of the population are presented in Table 6.

Table 6 Demographics of the patients

	Amsterdam UMC	MKA Kennemer & Meer	Radboudumc	Total
Number of patients	6	9	10	25
Gender (M/F)	3/3	6/3	3/7	12/13
Mean age at surgery (years)	39	27	29	31

2.3.4.1 Validation of OGA 2.0

The inter-observer and intra-observer intraclass correlation coefficients (ICC) and the mean differences for the maxilla, mandible, chin, left and right ramus are reported in Table 7-10 respectively.

Table 7 Intra-observer and inter-observer intraclass correlation coefficients and mean differences for measurements of the maxilla.

			Inter-observer		Intra-observer	
			ICC	Mean difference (±SD)	ICC	Mean difference (±SD)
Translation (mm)	RL	Center 1	0.996	0.064 (±0.062)	0.992	0.103 (±0.054)
		Center 2	0.999	0.045 (±0.042)	0.997	0.082 (± 0.050)
		Center 3	0.992	0.055 (±0.029)	0.992	0.048 (± 0.041)
		Mean	**0.996**	**0.055 (±0.042)**	**0.994**	**0.074 (±0.052)**
	AP	Center 1	0.938	0.251 (±0.245)	0.991	0.109 (± 0.077)
		Center 2	0.996	0.146 (±0.119)	0.997	0.097 (± 0.132)
		Center 3	0.991	0.089 (±0.113)	0.996	0.073 (± 0.059)
		Mean	**0.975**	**0.147 (±0.165)**	**0.995**	**0.085 (±0.091)**
	SI	Center 1	0.880	0.290 (±0.320)	0.973	0.168 (± 0.114)
		Center 2	0.934	0.289 (±0.310)	0.960	0.242 (± 0.261)
		Center 3	0.954	0.197 (±0.166)	0.961	0.176 (± 0.166)
		Mean	**0.923**	**0.261 (±0.259)**	**0.965**	**0.202 (±0.195)**
Rotation (degrees)	Roll	Center 1	0.988	0.076 (±0.109)	0.990	0.105 (± 0.052)
		Center 2	0.963	0.168 (±0.176)	0.961	0.174 (±0.186)

		Center 3	0.945	0.175 (±0.088)	0.960	0.140 (±0.096)
		Mean	**0.965**	**0.152 (±0.134)**	**0.970**	**0.144 (±0.130)**
	Pitch	Center 1	0.984	0.583 (±0.643)	0.997	0.332 (±0.111)
		Center 2	0.914	0.584 (±0.815)	0.904	0.526 (±0.860)
		Center 3	0.975	0.259 (±0.273)	0.972	0.275 (±0.273)
		Mean	**0.958**	**0.460 (±0.614)**	**0.958**	**0.378 (±0.550)**
	Yaw	Center 1	0.986	0.122* (±0.043)	0.988	0.092 (±0.059)
		Center 2	0.998	0.066* (±0.035)	0.997	0.066 (±0.070)
		Center 3	0.994	0.076 (±0.273)	0.997	0.055 (±0.030)
		Mean	**0.993**	**0.084 (±0.047)**	**0.994**	**0.065 (±0.054)**

SD = standard deviation, RL = right-left, AP = anteroposterior, SI = superior-inferior. Center 1 = Amsterdam UMC, Center 2 = MKA Kennemer & Meer, Center 3 = Radboudumc, * = statistically significant difference.

Table 8 Intra-observer and inter-observer differences and intraclass correlation coefficients for measurements of the mandible.

			Inter-observer		Intra-observer	
			ICC	Mean difference (±SD)	ICC	Mean difference (±SD)
Translation (mm)	RL	Center 1	0.997	0.110 (±0.095)	0.997	0.114 (±0.059)
		Center 2	0.995	0.099 (±0.085)	0.997	0.075 (±0.064)
		Center 3	0.993	0.115 (±0.075)	0.998	0.062 (±0.039)
		Mean	**0.995**	**0.107 (±0.082)**	**0.997**	**0.078 (±0.056)**
	AP	Center 1	0.995	0.237 (±0.201)	0.999	0.124 (±0.077)
		Center 2	1.00	0.081 (±0.064)	1.00	0.057 (±0.044)
		Center 3	0.993	0.147 (±0.129)	0.997	0.108 (±0.074)
		Mean	**0.996**	**0.147 (±0.143)**	**0.999**	**0.091 (±0.070)**
	SI	Center 1	0.999	0.242 (±0.194)	0.999	0.270 (±0.246)
		Center 2	0.999	0.116 (±0.097)	0.997	0.173 (±0.140)
		Center 3	0.983	0.311 (±0.272)	0.996	0.147 (±0.124)
		Mean	**0.994**	**0.226 (±0.220)**	**0.997**	**0.192 (±0.166)**
Rotation (degrees)	Roll	Center 1	0.981	0.163 (±0.151)	0.992	0.105 (±0.074)
		Center 2	0.929	0.330 (±0.557)	0.997	0.088 (±0.076)
		Center 3	0.943	0.129 (±0.118)	0.955	0.166 (±0.118)
		Mean	**0.951**	**0.228 (±0.352)**	**0.981**	**0.122 (±0.100)**
	Pitch	Center 1	0.993	0.638 (±0.543)	0.994	0.570 (±0.588)
		Center 2	0.994	0.266 (±0.299)	0.996	0.242 (±0.184)
		Center 3	0.975	0.475 (±0.260)	0.990	0.303 (±0.121)
		Mean	**0.987**	**0.392 (±0.293)**	**0.993**	**0.343 (±0.332)**
	Yaw	Center 1	0.993	0.142 (±0.228)	1.00	0.031 (±0.026)
		Center 2	0.997	0.103 (±0.137)	0.999	0.068 (±0.046)
		Center 3	0.988	0.129 (±0.064)	0.995	0.081 (±0.039)
		Mean	**0.993**	**0.126 (±0.140)**	**0.998**	**0.066 (±0.042)**

SD = standard deviation, RL = right-left, AP = anteroposterior, SI = superior-inferior. Center 1 = Amsterdam UMC, Center 2 = MKA Kennemer & Meer, Center 3 = Radboudumc.

Table 9 Intra-observer and inter-observer differences and intraclass correlation coefficients for measurements of the chin.

	Inter-observer		Intra-observer	
	ICC	Mean difference (±SD)	ICC	Mean difference (±SD)

Translation (mm)	RL	Center 1	0.986	0.230 (±0.231)	0.999	0.083 (±0.043)
		Center 2	0.968	0.257 (±0.281)	0.921	0.296 (±0.478)
		Center 3	0.933	0.168 (±0.113)	0.979	0.095 (±0.050)
		Mean	**0.962**	**0.223 (±0.210)**	**0.966**	**0.169 (±0.301)**
	AP	Center 1	0.974	0.197 (±0.215)	0.955	0.274 (±0.268)
		Center 2	0.958	0.238 (±0.219)	0.994	0.160 (±0.110)
		Center 3	0.981	0.240 (±0.108)	0.993	0.113 (±0.093)
		Mean	**0.971**	**0.213 (±0.158)**	**0.981**	**0.150 (±0.137)**
	SI	Center 1	0.980	0.299 (±0.177)	0.991	0.237 (±0.101)
		Center 2	0.968	0.144 (±0.218)	0.970	0.170 (±0.184)
		Center 3	0.979	0.301 (±0.285)	0.998	0.107 (±0.087)
		Mean	**0.976**	**0.251 (±0.245)**	**0.986**	**0.160 (±0.140)**
Rotation (degrees)	Roll	Center 1	0.980	0.199 (±0.252)	0.995	0.112 (±0.102)
		Center 2	0.967	0.415 (±0.611)	0.962	0.468 (±0.684)
		Center 3	0.99	0.213 (±0.187)	0.995	0.154 (±0.134)
		Mean	**0.979**	**0.285 (±0.410)**	**0.984**	**0.267 (±0.444)**
	Pitch	Center 1	0.938	1.057 (±0.629)	0.941	0.907 (±0.638)
		Center 2	0.898	0.681 (±1.225)	0.825	0.858 (±1.669)
		Center 3	0.995	0.505 (±0.442)	0.999	0.250 (±0.154)
		Mean	**0.944**	**0.654 (±0.824)**	**0.922**	**0.604 (±1.075)**
	Yaw	Center 1	0.994	0.395 (±0.343)	0.999	0.135 (±0.106)
		Center 2	0.982	0.348 (±0.512)	0.971	0.411 (±0.658)
		Center 3	0.968	0.311 (±0.182)	0.986	0.209 (±0.105)
		Mean	**0.981**	**0.345 (±0.362)**	**0.985**	**0.264 (±0.414)**

SD = standard deviation, RL = right-left, AP = anteroposterior, SI = superior-inferior. Center 1 = Amsterdam UMC, Center 2 = MKA Kennemer & Meer, Center 3 = Radboudumc.

Table 10 Intra-observer and inter-observer differences and intraclass correlation coefficients for measurements of the left and right ramus.

			Inter-observer		Intra-observer	
			ICC	Mean difference (±SD)	ICC	Mean difference (±SD)
Left ramus Rotation (degrees)	Auto	Center 1	0.947	0.954 (±0.830)	0.986	0.599 (±0.420)
		Center 2	0.952	0.331 (±0.269)	0.971	0.410 (±0.248)
		Center 3	0.953	0.781 (±0.759)	0.915	0.681 (±0.728)
		Mean	**0.951**	**0.673 (±0.684)**	**0.957**	**0.587 (±0.518)**
	Flare	Center 1	0.986	0.480 (±0.746)	0.999	0.177 (±0.089)
		Center 2	0.993	0.333 (±0.212)	0.997	0.243 (±0.252)
		Center 3	0.998	0.388 (±0.320)	0.996	0.258 (±0.285)
		Mean	**0.992**	**0.398 (±0.423)**	**0.997**	**0.239 (±0.236)**
	Roll	Center 1	0.951	0.333 (±0.327)	0.986	0.250 (±0.187)
		Center 2	0.997	0.100 (±0.000)	0.998	0.111 (±0.078)
		Center 3	0.992	0.300 (±0.262)	0.994	0.190 (±0.173)
		Mean	**0.980**	**0.208 (±0.204)**	**0.993**	**0.183 (±0.152)**
Right ramus Rotation (degrees)	Auto	Center 1	0.984	1.006* (±0.828)	0.992	0.413 (±0.558)
		Center 2	0.946	0.318* (±0.242)	0.961	0.359 (±0.347)
		Center 3	0.929	0.543 (±0.402)	0.972	0.269 (±0.188)
		Mean	**0.953**	**0.538 (±0.524)**	**0.975**	**0.348 (±0.351)**
	Flare	Center 1	0.993	0.560 (±0.464)	0.999	0.263 (±0.176)

		Center 2	0.866	0.886 (±0.795)	0.994	0.185 (±0.188)
		Center 3	0.989	0.450 (±0.372)	0.992	0.330 (±0.203)
		Mean	**0.949**	**0.622 (±0.599)**	**0.995**	**0.271 (±0.193)**
	Roll	Center 1	0.999	0.250 (±0.243)	0.998	0.167 (±0.225)
		Center 2	0.935	0.422 (±0.427)	0.994	0.144 (±0.113)
		Center 3	0.981	0.260 (±0.207)	0.996	0.110 (±0.099)
		Mean	**0.972**	**0.308 (±0.313)**	**0.996**	**0.142 (±0.135)**

SD = standard deviation, RL = right-left, AP = anteroposterior, SI = superior-inferior. Center 1 = Amsterdam UMC, Center 2 = MKA Kennemer & Meer, Center 3 = Radboudumc, * = statistically significant difference.

The mean inter-observer and intra-observer translational and rotational differences of the maxilla and mandible were all below 0.3 mm and 0.5 degrees. The least observer dependent was the anteroposterior translation of the mandible, for which an inter-observer and intra-observer ICC of 0.996 and 0.999, respectively, are reported in Table 4. The differences between the centers were non-significant, except for the interobserver difference of the yaw of the maxilla ($p = 0.047$) and the intraobserver difference of the autorotation of the right ramus ($p = 0.046$). These differences where however below 0.7 degrees and were therefore considered clinically insignificant.

Table 5 provides the results of the chin analysis. Concerning the translational differences of the chin, the superior-inferior direction was slightly more user dependent than the anteroposterior and right-left directions (0.251 mm versus 0.213 mm and 0.223 mm, respectively). The highest difference between users was reported in the pitch with 0.654 degrees inter-observer and 0.604 degrees mean intra-observer difference.

With regard to the left and right ramus, the autorotation of the left ramus and the flare of the right ramus were reported to be most user dependent (with maximal errors of 0.673 degrees and 0.622 degrees). Also, the reported inter- and intra-observer ICCs were all above the 0.94.

2.3.5 Discussion

The OGA 2.0 presented in the current study is a successor of the OGA presented in an earlier study [15]. Drawbacks of the previous OGA version were the absence of the possibility to assess the postoperative accuracy of the osseous chin, the dependence on a specific virtual planning software and the need for SBM for accurate matching of the rami. In the newly presented tool, the postoperative accuracy of the rami is assessed using VBM instead of SBM, the postoperative accuracy of the osseous chin can be assessed and the OGA 2.0 is no longer dependent on any planning software and can be used as a stand-alone program. The OrthoGnathicAnalyser 2.0 is developed to objectively quantify the movements of the individual segments of orthognathic surgery.

2.3.5.1 Validation results

The results of this multicenter validation study demonstrated a good reproducibility of the calculated results, with a maximum translational error of 0.26 mm and rotational error of 0.67 degrees, and corresponding high ICCs (>0.92). The current results of the maxilla were comparable to the results described in literature, with an inter-observer and intra-observer ICC of >0.97 and >0.98, for translation and rotation respectively [15, 28]. Stokbro and Thygesen used VBM for measuring the movements of the maxilla and found high ICC values similar to this current study [29]. The translational and rotational results of the mandible showed excellent reproducibility (ICC>0.99 and ICC>0.95 respectively) and were also

comparable to previous results [15]. The OGA 2.0 is a robust tool as minimal differences between the centers, and thus different manufacturers of scanners, were reported.

For the matching of the chin, preliminary tests were executed to evaluate which registration technique would perform best. During these tests it was observed that voxel-based matching resulted in less accurate alignment in the sagittal plane due to a deviation in the pitch. It was hypothesized that the result of the voxel-based matching algorithm was affected by the combination of the relatively small volume of the chin and the high-density fixation material. For this reason, it was chosen to implement SBM instead of VBM. This has resulted in a reproducible evaluation of the deviations of the chin segment, with low intra-observer and inter-observer differences (below 0.25 mm or 0.7 degrees). As these results for the analysis of the osseous chin are clinically acceptable it is worth noting that the inter-observer difference for the chin is systematically higher than the maxilla, mandible and rami. Underlying reason for this higher inter-observer difference could be the SBM, which required more input of the user.

2.3.5.2 Advantages current method

In our previous study [15], the matching of the left and right ramus was performed with SBM to counteract the image artifacts as a result of the sagittal split osteotomy. This technique has resulted in observer differences of more than one degree. Because of the reported difference and the described user dependency in the literature [17], the matching technique of the rami was changed to voxel-based matching in OGA 2.0 as there was an updated version of the voxel-based algorithm available. Without correcting the aforementioned image artifacts, the reported maximum error was almost halved to 0.6 degrees. Using VBM instead of SBM is more time efficient as the input from the user is minimized.

In the previously validated OGA, three landmarks for each jaw segment were required to construct a virtual triangle to allow for the calculation of the clinically relevant translational and rotational movements. Multiple landmark identification has been eliminated by voxel-based registration of the jaw segments. In the new version of OGA, a total of four landmarks needs to be identified instead of the twelve (three for each segment) in the earlier OGA version. Identification of only these four landmarks still enables the computation of the required calculations. As a consequence, the workflow becomes more efficient and further eliminates the inaccuracies as a result of multiple landmark identification [30].

2.3.5.3 Study limitations

The error caused by identification of the landmarks ranged from 0.02 to 2.47 mm [30-32]. Ideally, the manual identification step would be completely eliminated in the software. A promising development is the automatic 3D landmarking using artificial intelligence. Some recent studies have reported errors below 2 mm [33, 34], making automatic 3D landmarking a potential alternative. However, as the landmarks are not used for matching but only function as rotation points, the identification of the landmarks has become of little concern. The high ICCs and low intra- and inter-observer variations support this statement.

The results indicated that the pitch of the chin was still relatively more user dependent than the other variables. It should be explored whether the voxel-based matching method could be adapted to facilitate selection of greyscale values (i.e. selection of the upper threshold) or reorientation of the ROI box to enable exclusion of the high-density fixation material.

For the assessment of the accuracy of the mandible it is important that the postoperative (CB)CT scan was acquired in the optimal occlusion, to mimic the planned position. For this retrospective study, some scans were acquired in a suboptimal occlusion, which led to an overestimation of the discrepancy in the planned and postoperative outcome. Since the main goal of this study was to validate the novel software, it was chosen not to analyze the surgical outcomes and focus on the validation of the software. For any clinical study, it is imperative to provide proper instruction to the patient before the postoperative scan in order to be able to accurately assess the surgical outcome of the mandible.

2.3.6 Conclusions

In conclusion, the reported results of this study demonstrated an excellent reproducibility (ICC of >0.92) of the quantification of the skeletal movements between two (CB)CT sets by the OrthoGnathicAnalyser 2.0. By implementing the chin analysis in this software tool, all surgical bony segments can be objectively evaluated and compared to the preoperative virtual plan. The OrthoGnathicAnalyser 2.0 allows an increased number of evaluations of orthognathic procedures.

3. Multicenter intervention study

3.1 Introduction

After validation of the implemented chin analysis in OrthoGnathicAnalyser 2.0, the goal was to investigate whether the patient-specific guidance system affects the accuracy of the execution of the genioplasty. The gold standard to prove a difference in treatment effect is a randomized controlled trial [35]. Compared to the validation study, this study type requires more extensive preparations. Within the available time and given circumstances, the study could only be prepared (not executed) during this research project. The next paragraph will contain an extensive 'Materials and Methods' section, to explain the proposed study and to illustrate the steps undertaken to initiate a prospective interventional study. The last paragraph of this chapter contains a personal reflection on the process and future steps concerning the multi-center intervention study.

3.2 Materials and methods

3.2.1 Ethical approval

Before starting a study concerning humans it must be considered whether the research is subject to the Dutch Medical Research Involving Human Subjects Act (WMO). This is the case if the following criteria are met [36]:

1. It concerns medical scientific research and
2. Participants are subject to procedures or are required to follow rules of behavior

For the intended multicenter randomized intervention study, both criteria are met. Thus, the study protocol must be approved by an accredited medical ethics committee (in Dutch abbreviated to METC). It was chosen to submit a request for approval to the METC of Amsterdam UMC, location AMC. A review by the METC requires the submission of a research file. This file needs to contain all the documents describing different aspects of the study. The required documents differ per type of study. The proposed intervention study concerns medical devices, which are custom-made. The medical devices in question have a risk classification of IIA, since it is surgically invasive and transient in use (rule 6 in [37]). Since the medical device is custom-made, no CE-certificate is required, but instructions for cleaning and sterilization are obligatory. In Table 11, the complete list of required documents for this study is provided.

Table 11 Content of research file required for the ethical review

Code	Document title
A1	Cover letter to reviewing committee
B1	ABR Form
B2	AMC Appendix
C1	Research protocol
D4	Instructions for cleaning and sterilization of medical devices
E1	Information letter research subjects
E2	Consent form research subjects
H1	CV independent expert
H2	CV coordinating investigator
I1	List of participating centers
I2	Research declarations

I3	CVs principal investigators
K3	Clinical Trial Agreement
K6	Confirmation application central monitoring Data Protection Impact Assessment Risk Assessment Data Management Plan

After the first submission, the METC has discussed the research project during a meeting and presented some feedback and/or questions on the proposal. This has led to some minor changes and clarifications in the protocol. After a last round of feedback, approval was obtained for the proposed study.

3.2.2 Patients

3.2.2.1 In- and exclusion criteria

Base: Orthognathic surgery patients (>18 years) in whom a 3D preoperative planning is made and in whom a genioplasty is indicated.

Inclusion criteria: In order to be eligible to participate in this study, a subject must meet all of the following criteria:

- Older than 18 years
- Genioplasty indicated in 3D preoperative planning.

Exclusion criteria: A potential subject will be excluded from participation in this study when the surgery is indicated for any of the following conditions:

- Congenital disorders (e.g. craniofacial microsomia)
- Obstructive Sleep Apnea Syndrome (OSAS)
- Transgender surgery
- Previous orthognathic surgery

3.2.2.2 Power analysis

In the proposed intervention study, there will be six primary outcome measures: translation in mediolateral direction, translation in anteroposterior direction, translation in superior-inferior direction, rotation in the frontal plane (roll), rotation in the sagittal plane (pitch) and rotation in the axial plane (yaw). The absolute difference between the planned value for a parameter and the postoperative acquired value will be calculated (e.g. 1 mm translation to the left planned and 2.2 mm acquired postoperatively gives a difference of 1.2 mm). In literature, one study has been performed on the effect of genioplasty guides on surgical result [12]. They did report differences between planned and acquired position for each of the six outcome measures, but the reported values were not absolute. Since the amount of error is of interest rather than the direction of the error, using the absolute value is highly recommended.

Access to the raw data of the aforementioned study was not granted by the authors after a request (via e-mail to the corresponding author), so the mean and standard deviations as reported in the article were used (Table 13) for a power analysis. To enable a power analysis for the proposed study parameters, the reported values needed to be transformed to absolute values. A folded normal distribution would result from the data transformation, assuming a normal distribution of the reported outcome parameters. The folded mean and folded standard deviation that describe the folded normal distribution were acquired using Table 1 in [38]. The obtained

values are provided in Table 13. These values used to calculate the effect size for the power analysis.

Table 12 Mean and standard deviation of the error values reported by Li et al. and of the folded normal distribution

	Li et al.		Folded normal distribution	
	Experimental	*Control*	*Experimental*	*Control*
	Mean ± std	*Mean ± std*	*Mean ± std*	*Mean ± std*
Mediolateral [mm]	-0.01 ± 0.62	0.04 ± 1.47	0.48 ± 0.39	1.17 ± 0.89
Anteroposterior [mm]	-0.84 ± 0.74	-1.22 ± 1.76	0.94 ± 0.62	1.73 ± 1.27
Superior-inferior [mm]	0.02 ± 0.85	0.60 ± 2.59	0.68 ± 0.51	2.14 ± 1.61
Roll [degrees]	0.33 ± 1.76	-0.60 ± 2.69	1.45 ± 1.09	2.21 ± 1.67
Pitch [degrees]	-1.51 ± 2.17	-4.00 ± 6.06	2.14 ± 1.56	5.86 ± 4.30
Yaw [degrees]	-0.04 ± 1.53	-0.70 ± 2.87	1.22 ± 0.92	2.36 ± 1.78

Std = standard deviation

For the a priori sample size calculation, the software G*power (Heinrich Heine Universität Düsseldorf, Germany) was utilized. Since the absolute values are used, normality cannot be assumed. A Wilcoxon-Mann-Whitney test (two groups) was chosen for this reason. Since six outcome parameters are used, a Bonferroni correction was applied to counteract the effect of multiple comparisons. The following values were entered: alpha of 0.0083 (=0,05/6) and power of 0.8. The resulting sample sizes per group are reported in Table 14.

Table 13 Sample size per group specified per study parameter

	Mediolateral	Anteroposterior	Superior-inferior	Roll	Pitch	Yaw
Sample size per group	28	43	20	90	22	42

The sample size of 90, found for the roll parameter, would be infeasible in clinical practice. It was chosen to accept the possibility of not proving a statistically significant difference in the roll variable. It was decided to include 43 patients per group (total of 86 patients), similar to the number of patients analyzed in the study by Li et al.

3.2.3 Study design

Primary Objective: to determine if the application of patient-specific guidance system improves the accuracy of the genioplasty.

Secondary Objective: to determine if the application of patient-specific guidance system affects the surgical time.

Type of study: Prospective multicenter randomized intervention study

Study groups: Experimental and conventional group

- Experimental group: genioplasty will be executed with a 3D-printed cutting and repositioning guidance system. This will be designed based on the patient's anatomy, the desired osteotomy and the desired position of the chin segment. Examples of the guides are depicted in Figure 10.
- Control group: genioplasty will be executed without patient-specific guidance system.

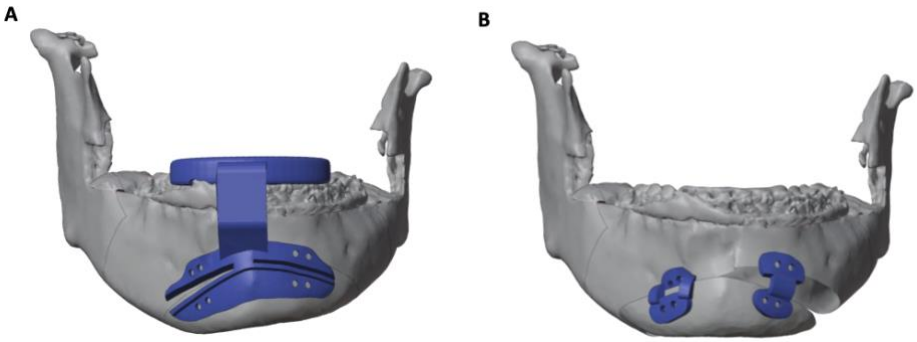


Figure 10 Intended design of the patient-specific guidance system. A) The cutting guide (blue) is supported by the dentition and bone (grey). Screw hole are drilled and used for fixation during the osteotomy. B) The repositioning guides (blue) are positioned using the screw holes from the cutting guide.

Setting of the study: For this study, a multicenter approach is chosen to ensure sufficient patient inclusion within the duration of the study. Participating centers will be the Amsterdam UMC location AMC, Radboudumc and MKA Kennemer & Meer location Haarlem.

Overview of study procedures: Figure 11 visualizes the process of the current clinical practice for standard orthognathic surgical patients (blue) and the required additions to the clinical protocol necessary for this study (green). If no data is collected for the study, the boxes are grey. Study inclusion will start at preoperative setup (t = 2) and participation will end after 6 weeks of follow-up with a CBCT-scan and analysis (t = 5). While designing the individual steps of the study, care was taken to minimize the additional burden to the patient.

Patients are referred to the maxillofacial surgeon by an orthodontist before the start of any treatment. Patients visit the surgeon to gain informed consent about a combined orthodontic and surgical plan (t = 0). The goal of this appointment is to determine if the patient wants to undergo a combined orthodontic and surgical treatment. If this is the case, the patient will be referred back to the orthodontist to start orthodontic pre-treatment. When this orthodontic pre-treatment is finished, the patient will come back for the 'ready for surgery' appointment (t = 1). During this appointment, the surgeon will evaluate if the orthodontic preparation is sufficient to perform the surgery. If so, the patient will be put on the waiting list for surgery. All aspects of surgery are discussed a second time during this visit.

The intended moment for patient recruitment is the set-up appointment (t = 2), which takes place ± 4 weeks before the surgery. The goal of this appointment is to perform detailed examinations of the patient and to acquire a (CB)CT scan and dental impressions necessary for virtual surgical planning, which is current clinical protocol. If a genioplasty might be necessary, based on clinical judgement, the surgeon will ask a member of the research team to inform the patient about the study. The patient will be given a patient information folder and an informed consent form.

The 3D preoperative planning is performed two weeks before the surgery for every orthognathic patient (t = 3). During the 3D evaluation, the indication for a genioplasty has to be confirmed (clinical protocol). If a genioplasty is indeed required, a member of the research team contacts the patient to confirm the inclusion. If the patient is willing to participate, the informed consent form (which is already in the possession of the patient) will be signed by the patient and sent back to the surgeon by mail. Randomization into the experimental (with guidance system) or control (conventional) group will take place. If the patient is assigned to the experimental group, a patient-specific cutting and repositioning guide will be designed and 3D-printed together with the patient-specific splints for maxillary and/or mandibular positioning. If the patient is assigned to the

control group, only the patient-specific splints for maxillary and/or mandibular positioning will be printed. At admission before the surgery (t = 4), the informed consent can be checked a second time.

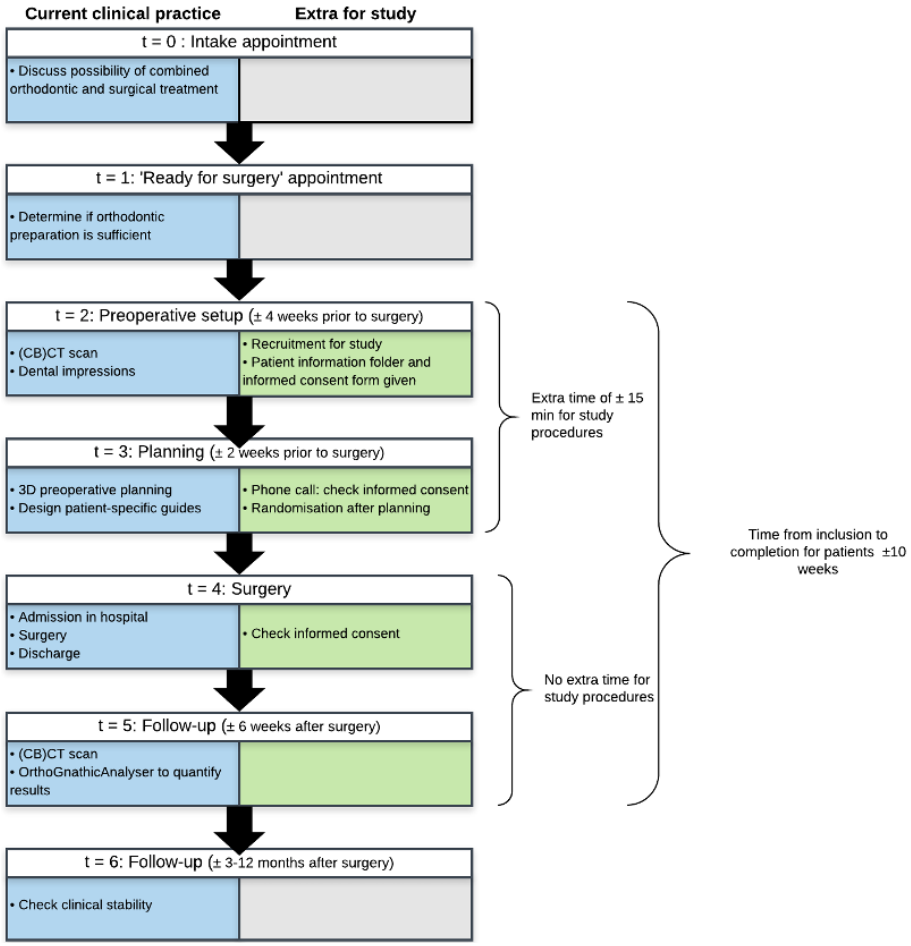


Figure 11 Process of current clinical practice for orthognathic surgeries.

During the 6-week follow-up appointment (t = 5), it is current clinical practice to perform a (CB)CT scan to enable analysis of the result of the orthognathic surgery by means of the OrthoGnathicAnalyser software [15]. The difference between the planned and acquired position of all the segments (including the genioplasty) are shown in this software. The differences are expressed in rotational parameters and translational parameters in all three dimensions (roll, pitch, and yaw in degrees and mediolateral, anteroposterior and superior-inferior translation in millimeters). These are the six primary outcome measures. During the 3 months postoperative follow-up appointment (t = 6), the clinical stability will be evaluated.

All study data is collected from the preoperative setup appointment (t = 2) until the 6-week follow-up appointment (t = 5). It is estimated that this will take 15 minutes of extra time, which is mostly used to inform the patient about the study, to gain informed consent (10 min) and to confirm study participation with the patient after the planning phase (5 min, phone call). At other time points, the current clinical protocol is followed. No data is gathered at subsequent follow-up after completion of the study at the six weeks follow-up.

3.2.4 Randomization

To establish that an intervention yields a certain clinical outcome, patients need to be randomly allocated into an intervention and control group [39]. The process of randomization ensures that each patient has an equal chance of receiving one of the two treatments, independent of patient characteristics [40]. It also ensures that the two groups being compared are similar in both measured and unmeasured patient characteristics. This is necessary to guarantee an unbiased result [39].

Pure random allocation may lead to imbalances in group sizes [41]. This imbalance can be minimized by applying restriction techniques. The most common and well-established method is the permuted block design (PBD) [39]. Patients are randomly assigned to a treatment group within a certain set of study participants, also called a block. Within the block, assignment to a treatment is performed in a random order, while maintaining the desired allocation ratios.

A basic method for assigning treatment within the blocks is random number generation. In this method, a number is randomly generated for each assignment. For example, random numbers are generated for a block of six patients with treatment groups intervention (I) and control (C), as depicted in Table 14. Subsequently, these numbers are ranked from highest to lowest to determine the allocation order. In the example in Table 15, this results in the allocation order CICIIC for the first block.

Table 14 Permuted Block Design randomization explained.

Randomly generated numbers		Ranked numbers	
Treatment group	*Random number*	*Treatment group*	*Random number*
Intervention	2	Control	56
Intervention	14	Intervention	35
Intervention	35	Control	20
Control	20	Intervention	14
Control	1	Intervention	2
Control	56	Control	1

Left column: Random numbers ranging from 0 to 100 are generated for the block of six patients with treatment groups intervention (I) and control (C).
Right column: After ranking the randomly generated numbers from highest to lowest an allocation order arises, which is in this case: CICIIC

As illustrated above, implementation of PBD can be quite simple. Also, PBD helps to avoid chronological bias, as the allocation is performed independent of the time. However, PBD is prone to selection bias [42]. Considering the deterministic nature of the last allocation of a block, selection bias will occur if the block sizes are known. For the example above, after the assigned CICII, it is certain that the last patient will be assigned to the C group. A solution to this problem would be allocation concealment (blinding the researcher). This however is not possible in the proposed study, considering the obvious difference in treatment protocol. Another solution is the application of random block sizes, in which the researcher does not know which patient will be the last of a block and cannot predict the treatment allocation [42].

Since the proposed study will be multicenter, it must be ensured the allocation between treatment groups within the participating centers is also balanced. Therefore, the permuted blocks are stratified for each center. This ensures that a discontinuation of the participation of a center or poor enrolment will not affect the overall balance of the treatment groups [39].

3.2.5 Data collection

From all the documents in Table 12, the Data Protection Impact Assessment (DPIA) is most concerned with data collection. This is a document which has been introduced to ensure the study will comply to the General Data Protection Regulation (GDPR). The DPIA forces the researcher to think beforehand which data collection software will be used, which data will be collected and how the data will be transferred between centers.

The other important document concerning data is the Data Management Plan (DMP). The DMP is used to document how the data is managed and stored. This document can be updated during the study. The following sub-paragraphs will describe the data collection software, data to be collected, data flows and data management.

3.2.5.1 Data collection software

While completing the DPIA, it became apparent that the data collection required more sophisticated software than SPSS or Excel. The recommended alternative is Castor Electronic Data Capture (EDC) (Castor, Amsterdam, the Netherlands). Castor EDC is a cloud-based data solution which allows secure database building in multiple centers. It is GCP and ISO27001 certified and recommended for collecting data in investigator-initiated research. Since the software is specifically designed for data collection in medical trials, it facilitates several important features, such as an audit trial, authorizations per user, randomization and lockdown of the database after the end of treatment of last participant.

In order to be able to build a study in Castor, a beginner's course, offered by the Amsterdam UMC, was followed. This course explained the basic knowledge essential to build your own study. Once the study is fully implemented in Castor, the study status can be set to live. Whenever a patient is included in the study, the researcher can create a new 'record' (or patient) and start completing the electronic Case Report Forms (eCRFs) per patient. After ensuring the patient meets the inclusion criteria, the patient will be automatically randomized in the experimental or control group, with variable block randomization stratified per center.

3.2.5.2 Data to be collected

For each patient, the research team will collect the following data from patient specific data software (Epic) and store it in Castor:

- *Demographic data*: information such as age at time of surgery, gender, indication, type of surgery and type of genioplasty. These variables will be used to describe the research population and compare between the two groups.
- *OrthoGnathicAnalyser data*: the six variables describing the translational and rotational differences of the chin segment. These will be used for the primary goal.
- *Surgical data*: information such as operator experience, whether the surgery was according to plan, applied technique in the conventional group, fixation material and duration of the surgery. The latter variable will be used for the secondary goal, the others to explore possible bias.
- *Follow up data*: possible complications will be collected. This will be used for safety reporting of the medical device.

3.2.5.3 Data flows

The flowchart in Figure 12 illustrates the different information flows of the entire study. The current clinical processes are displayed in black, the additional processes for the study are displayed in red. t = # refers to the moments in Figure 11 in paragraph 3.2.3.

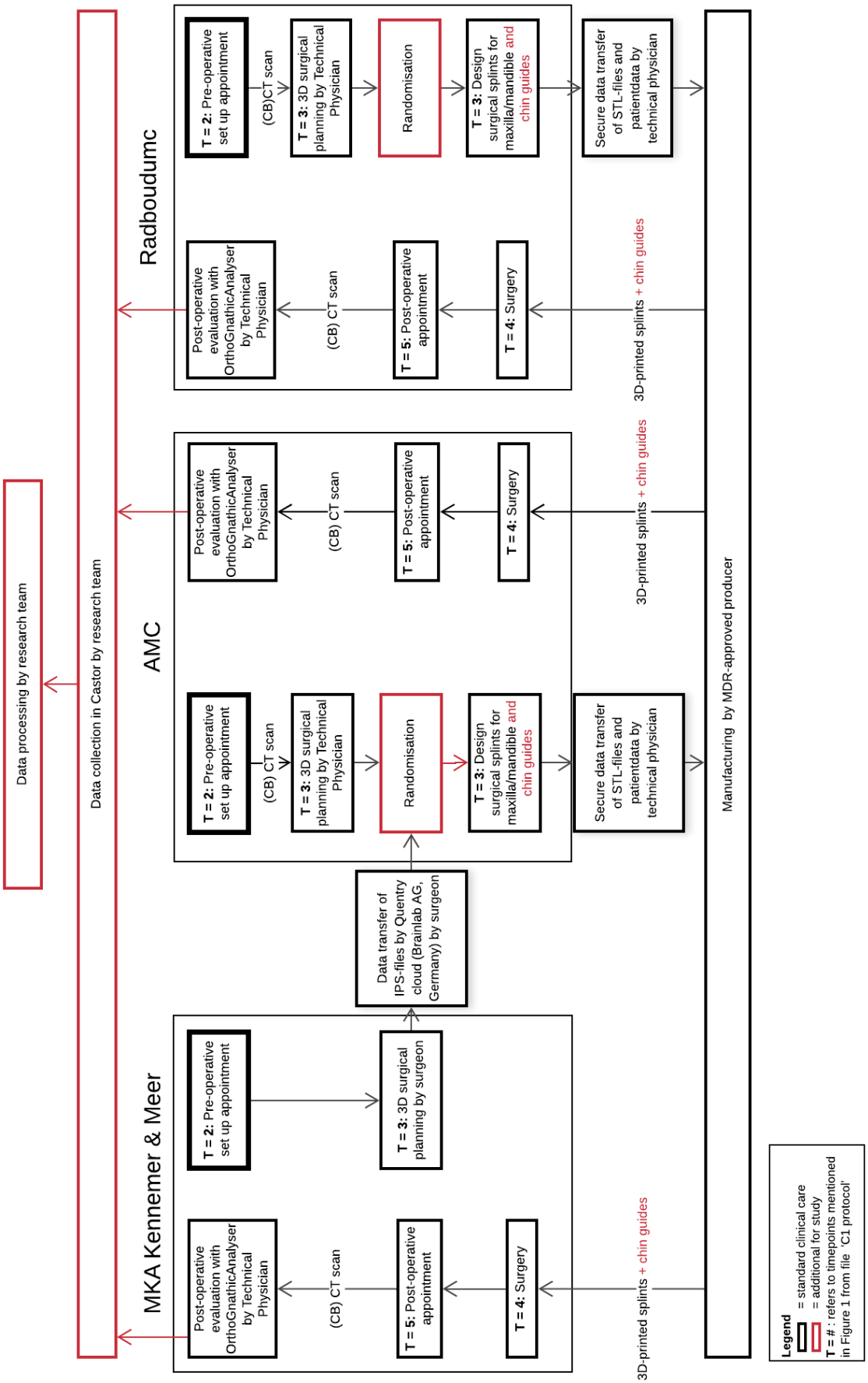


Figure 12 Information flows of the proposed study. The black boxes depict standard clinical care. The red boxes are additional for the study

The Amsterdam UMC, location AMC and Radboudumc are able to perform the virtual planning and design the guides independently. MKA Kennemer & Meer already is used to have the guides designed at Amsterdam UMC, location AMC and to send the virtual planning file (created with the planning software IPS CaseDesigner) to the AMC through Quentry Cloud (Brainlab AG, Munich, Germany) at t = 3. Subsequently, the technical physicians at the AMC design the required guides. This will also take place in the proposed study.

Currently, the 3D-models (as Standard Tessellation Language files (STL-files)) of the splints (to move the mandible and maxilla) and CT-images (as Digital Imaging and Communications in Medicine (DICOM) files) are securely sent to the manufacturer on a routine basis. In the context of this study, additional STL-files of the chin guides will be sent along with the 3D splint models. The manufacturer sends the 3D-printed splints and chin guides to the hospital in which the surgery (t = 4) will take place (through registered mail). For this study, the additional chin guides will be sent in the same package.

After the postoperative analysis is performed with OrthoGnathicAnalyser software, the result is saved as a PDF-file in the Amsterdam UMC patient specific data software (Epic). The researcher will acquire access to the local protected work environment of the participating hospitals. The data will be pseudonymized and entered in Castor. The subject identification log will be kept on the specific site, separate from the study database.

The research team in the AMC will centrally download the data from Castor and save it on the department's secure drive (G-drive). The data can then be used for statistical analysis in SPSS.

3.2.5.4 Data management

In an effort to keep track of all the study documents and data, the Amsterdam UMC has provided a template for a data management plan (DMP). The research team can complete this document during the course of the study. The document divides the study in five phases: 1) study preparation, 2) data collection, 3) processing & statistical analysis, 4) writing & publishing and 5) archiving & open data. In each phase crucial tasks within subjects are identified and can be checked off when performed. When the task results in a certain document, the specific path to the document needs to be specified. In this way, it can be used as a source document to keep track of what is stored where. For now, only the first phase could be completed. The following subjects were covered: privacy and security safeguards, data collection, data storage and data sharing.

3.2.6 Data analysis

3.2.6.1 Study parameters

Primary study parameters: There are six main study parameters in this study: three describing the absolute translation (anteroposterior, mediolateral and superior-inferior) and three describing the absolute rotation (pitch, roll and yaw). To compute these variables, the methodology of OGA will be used [15], since this method provides reproducible, quantitative outcome measures in three dimensions. The planned movement of the genial segment can be expressed in rotation (roll, pitch, yaw) and translation (mediolateral (x-dimension), anteroposterior (y-dimension), superior-inferior (z-dimension)) from its preoperative position. The postoperative (CB)CT scan can be superimposed on the preoperative virtual planning. Consequently, the preoperative genial segment can be superimposed on the postoperative position of the genial segment. The actual movement of the genial segment, in rotation and translation parameters, is automatically obtained

from the superimposition process. The absolute difference between planned and actual rotations and translations provides an intuitive accuracy measurement of the genioplasty. A difference between the control group and experiment group of more than 1 mm (translation) or 2 degrees (rotation) is regarded as a (clinical) significant difference and thus a positive outcome.

Secondary study parameter: The secondary outcome parameter is the surgical time. This will be defined as the time between the start of incision for the genioplasty and the fixation of the last screw. The surgeon will be asked to report this time after each surgery.

3.2.6.2 Statistical analysis

Primary study parameters: The six accuracy parameters of the genioplasty will be calculated by the OGA. Since the amount of error is of interest rather than the direction of the error, the absolute value of the computed differences will be taken. Due to this transformation, it is assumed the data will not be normally distributed. Thus, six Mann-Whitney U tests will be used to assess the significance of mean differences in all directions between experimental and control groups. Since six outcome parameters are used, a Bonferroni correction will be applied to counteract the effect of multiple comparisons. If data are missing for any of the variables, the subject will be excluded from the data analysis.

Secondary study parameter: The secondary study parameter is the surgical time. Since it is expected that the recording of this variable may be forgotten sometimes, missing this data point will not lead to exclusion of the patient from the study. This will be a quantitative continuous variable. Shapiro-Wilk test will be used to determine if the data is normally distributed. If so, an independent samples t-test will be performed to test for a statistically significant difference between groups; if not, a Mann-Whitney U test will be used.

3.3 Discussion

The preparation for the multicenter intervention study has provided both substantive and personal learning experiences. Substantively, hands-on experience on the preparation and acquisition of an ethical approval was gained. Since it was the first time preparing a research file for an METC and since there is no clear roadmap available that illustrates how this should be done, a lot of information gathering was required to determine what was necessary to acquire the permission. It came down to the preparation of several documents based on provided templates. Often an individual document referred to a certain expert that needed to be contacted to discuss its content. This has considerably lengthened the process. At first, the process felt a bit confusing and excessive. However, during the process things started to come together. The individual documents were referring to each other, completing a circle eventually. In the final document, every possible subject concerning the study was thought-through and well-documented. It was learned that the process of requesting METC approval forces researchers to approach the study design in a structured manner.

The personal learning experience was to (almost) independently manage and organize a project of this size. Also, different experts across the hospital needed to be consulted. This required perseverance and time management skills.

Due to the COVID-19 situation, the study was not allowed to start. Once the circumstances allow for continuation of clinical studies, a PhD student will take over the responsibilities of this research project. First and next steps will be setting up a monitoring plan and starting patient inclusion.

4. General discussion

This chapter discusses the main results obtained from this research project. The goal was to validate 3D-printed guidance systems for the execution of genioplasties with a new version of OrthoGnathicAnalyser (OGA). For this purpose, three different projects were undertaken; a literature review (see Appendix A), a validation study of the newly developed OrthoGnathicAnalyser 2.0 and preparations for a multicenter intervention study. These projects are discussed separately in the following paragraphs.

4.1 Literature review

A literature review was performed to put 3D-printed guidance systems for genioplasties in the context of orthognathic surgery and to give an overview of the current developments of the guidance systems. From this review, three important conclusions were drawn.

First, a 3D quantification tool was identified as a prerequisite to evaluate the surgical accuracy of the 3D-printed guidance system. Such a tool would allow objective and precise quantification of all possible translations and rotations of the chin segment. In the literature, a wide variety of different assessment methods of orthognathic surgery was reported. In an attempt to identify a universal assessment protocol, the review of Gaber et al. has compared seven different methods for the assessment of orthognathic surgery [14]. Their conclusion was that the absence of consensus between different centers impeded comparison of the results between different studies. The semi-automatic 3D tool, OGA published by Baan et al [15], appeared to be the method of choice, due to the application of voxel-based registration and the ability to quantify all possible translations and rotations. However, this tool was not capable of analyzing the surgical result of the chin. To prevent an excessive number of assessment methods, it was chosen to build upon the OGA and implement a dedicated chin analysis.

Second, evidence to support the need for a randomized controlled intervention study, comparing genioplasties executed with and without a 3D-printed guidance system, was found. Only the study of Li et al. approached the study design required to draw hard conclusions on the accuracy of the guidance system [12]. This study compared two groups (with and without guides) of considerable group size (44 patients in each arm) and computed translations and rotations in all three dimensions. Results indicated a superior surgical accuracy when surgical guides were used. However, this study was not randomized controlled, making the results sensitive to bias.

Thirdly, the 3D-printed guidance system which consisted of cutting and repositioning guides, as described by Li et al [12], was identified as most promising. This design was tested on a relatively large population and accurate results were reported. Positive experiences were gained during internal audits of this design in the clinical setting of Amsterdam UMC. Based on these arguments, it was chosen to apply this design in the intervention study.

4.2 Validation study

To reliably evaluate surgical results, it was deemed necessary to implement the analysis of the chin in the current software of Baan et al. [15]. In the subsequent multicenter validation study, the OrthoGnathicAnalyser (OGA) version 2.0 demonstrated excellent reproducibility (ICC of >0.92) of the quantification of the skeletal movements realized by the surgery. The

final version of the scientific paper will be submitted to the journal PLOS ONE (San Francisco, US).

4.2.1 Benefits

The most important modification in OGA 2.0 was the implementation of the chin analysis. While the tool was already capable of analyzing the surgical result of the mandible, maxilla and rami, it could not analyze the chin. As a genioplasty is part of the arsenal of orthognathic surgery to fully correct facial asymmetry, this implementation has enabled the evaluation of every facet of orthognathic surgery. The 3D assessment tool can therefore be applied in a variety of studies. For example, to evaluate the effect of 3D-printed guides for the genioplasty on the surgical accuracy, but also to evaluate the long-term effect of the repositioning of the rami caused by orthognathic surgery on condylar remodeling.

Advantageous of the OGA version 2.0 is the independence on any planning software. The required 3D models of the planned and preoperative segments can be produced in the planning software preferred by the user. Also, the validation study has proven that the software tool can handle image data acquired by scanners of different manufacturers. This implicates that the results of different studies can now be analyzed by this tool to enable comparisons between studies.

Compared to the previous version of OGA, the current version required less manual interaction (four landmarks need to be identified versus twelve), while still enabling the calculation of the clinically relevant translational and rotational movements. This has led to a further reduction of the inaccuracies as a result of multiple landmark identification [30]. In addition, the minimal user input saves time and allows for evaluation of the results executed by less experienced users.

4.2.2 Limitations

Manual identification of landmarks is still required in the current version of OGA. Ideally, this would be completely eliminated to minimize the identification error. A promising development is automatic 3D landmarking using artificial intelligence, for which errors below 2 mm are reported [33, 34].

For the matching of the chin, it was now chosen to implement SBM instead of VBM to ensure accurate alignment of the pitch. This decision was based on the hypothesis formulated after preliminary testing; the result of the VBM algorithm was affected by the combination of the relatively small volume of the chin and the high-density fixation material. The application of SBM has resulted in reproducible results, with a low intra-observer and inter-observer differences (below 0.25 mm and 0.7 degrees). However, the pitch was relatively more user dependent than the other variables. Therefore, it should be explored whether VBM could be adapted to facilitate the exclusion of the fixation material, either by selection of the greyscale values or by reorientation of the region of interest box.

4.3 Multicenter intervention study

To assess the effect of the 3D-printed guidance system on surgical accuracy, a multicenter randomized intervention study was initiated. Since this study concerns medical scientific research and the participants will be subjected to medical procedures, an extensive research proposal had to be prepared to gain ethical approval. This proposal had to cover different topics, such as the study design (paragraph 4.2.3) and data collection (paragraph 4.2.5).

After processing the feedback of the ethical committee, permission to start patient inclusion was obtained. The trial will start whenever the (COVID-19) circumstances allow for continuation of patient care and clinical studies.

4.3.1 Benefits

This multicenter randomized intervention study can provide strong evidence supporting the improved surgical accuracy of the genioplasty caused by the application of the 3D-printed guidance system. The multicenter approach ensures enough patient inclusion within the duration of the study. As different surgeons will perform the surgery, it will also be possible to evaluate the inter-surgeon variability. For the randomization it was chosen to implement a permuted block design with random group sizes, stratified per center. This will ensure an overall balance between the treatment groups and prevent any form of bias.

The secondary outcome parameter is the time required for the execution of the genioplasty during the surgery. Recording of this parameter will require minimal additional effort during the surgery and it can be indicative of the effect on the cost-benefit ratio. While no reports on this parameter were found in the literature, it would be interesting to know whether the application of the 3D-printed guides will increase or decrease the surgical time.

4.3.2 Limitations

The main goal of orthognathic surgery is to improve the function and esthetics of the patient. In the intended patient population, the chin will only affect the esthetics. Therefore, it could be argued that the actual efficiency of the intervention should be represented by the patient satisfaction. For this study, it was chosen to measure the effect on the surgical accuracy first, to enable isolation of the effect of the genioplasty from the effect of the bimaxillary surgery.

Another limitation is the potentially small improvement in the cost-benefit ratio. In the intervention study, benefit could be simulated with the surgical accuracy. While the chin segment is planned with submillimeter accuracy, the potential improvement could only be a few millimeters or degrees. On the other hand, the costs will be increased because of the additional 3D-printing and the design of the guides (2-4 hours by an experienced designer).

The last limitation is that the 3D-printed guidance system will be tested on a strictly defined research population, in the Netherlands. Exclusion criteria were applied to ensure a homogeneous research population. The excluded patient groups (patients with obstructive sleep apnea syndrome (OSAS) and congenital asymmetrical deformities) often require larger rotations and/or translations to fully correct the chin deformity. These difficult chin cases could potentially benefit more from the 3D-printed guides.

4.3.3 Future work

When the intervention study is completed and the results indicate that the accuracy of the genioplasty has improved, possibilities to improve the cost-benefit ratio could be explored. For example, feedback on the design gathered during the study could be used to improve the user-friendliness, aiming for a decrease in surgical time. An additional opportunity could be to automatize the design process, decreasing the required time of the designer and thus the costs.

In future studies, the final outcome parameter should be the patient satisfaction. The patient satisfaction can be measured by the Orthognathic Quality of Life Questionnaire (OQLQ) [43], which is specifically designed and validated for patients with severe dentofacial deformities. It

can be considered to include additional questions to emphasize the effect of the chin. Only when a superior accuracy due to the 3D-printed guides has been proven, this modified questionnaire will be used in future research. This can relate the increase of accuracy, caused by 3D-printed guides, to the effect on patient satisfaction. In this way, the effect of the bimaxillary surgery on the patient satisfaction will be eliminated.

The scope of the research could be expanded in future work. The acquisition of additional imaging one year after the surgery could help to analyze the long-term outcome. The comparison of the long-term to the short-term postoperative (CB)CT scan enables the evaluation of the stability of the genioplasty. It is hypothesized that advancements of the chin elongate the muscles which could increase the tension exerted on the chin segment, influencing the pitch and/or anteroposterior translation of the chin. Evaluation of the previously excluded patient populations must be explored. Application of the 3D-printed guides for the genioplasty patients, for instance, with OSAS. For these patients the direction of the osteotomy line is important, as it affects the muscle attachments, which in turn deliberately affect the position of the hyoid bone and the tongue [44].

5. Conclusions

To validate a 3D-printed guidance system for genioplasties, three different projects were executed during this research projects. First, a literature view was performed. From this review it was concluded that the scientific gap, supporting the use of the 3D-printed guides, endorsed the need for a prospective randomized controlled trial. The semi-automatic 3D assessment tool called OrthoGnathicAnalyser for the quantification of the maxilla and mandible was identified as most promising. This was based on the fact that it applied voxel-based matching, which is preferred because of its user independence, and that it computes the difference between the virtual plan and postoperative results for all possible translations and rotations per bony segment. Unfortunately, analysis of the chin was not implemented in this tool. Lastly, the most promising design of the 3D-printed guidance system was identified. The guidance system consists of a separate guide which is stabilized on the occlusal plane and which dictates the osteotomy. The other set of guides is used for the repositioning of the chin segment. This guidance system will be applied in the intervention study.

Subsequently, the OrthoGnathicAnalyser 2.0 has been developed. Most effort was put in the implementation of the analysis of the chin. Software development revealed an inferior accuracy of the pitch when voxel-based matching was applied for registration of the chin. While surface-based matching showed a better accuracy for the pitch and an overall good precision, it was chosen to implement this technique in the software. In addition, the chin analysis was based on the difference between the postoperative chin and the planned chin relative to the realized mandible instead of the planned mandible. This ensured that the results of the chin analysis would not be affected by a positioning error of the mandible.

After the software development, the OrthoGnathicAnalyser 2.0 was validated in a multicenter retrospective study for orthognathic surgery with genioplasties. For the chin in particular, a maximum inter-observer and intra-observer translational error of 0.25 mm (in superior-inferior direction) and rotation error of 0.65 degrees (of pitch) were reported. The error of the translation was comparable to the values of the mandible and maxilla, while the error of the rotation was slightly higher compared to the other segments. This entailed that the pitch is still relatively more user dependent. A possible improvement would be the implementation of voxel-based matching that facilitates the exclusion of the high-density fixation material. Overall, the reported results of this study demonstrated a good reproducibility (ICC of >0.92) of the quantification of the skeletal movements between two (CB)CT sets by the OrthoGnathicAnalyser 2.0. The validation of this 3D assessment tool is fundamental for comparing any (future) intervention in any patient group, as it is capable of objectively quantifying the surgical accuracy for each bony segment in orthognathic surgery.

A multicenter randomized controlled intervention study has been initiated to validate the 3D-printed guidance system for genioplasties. To acquire ethical approval by the local ethics committee, an extensive research file needed to be prepared. Approval to start the study has been acquired within this graduation project (METC 2020_036) and patient inclusion will start when Covid-19 circumstances allow for continuation of clinical studies and regular patient care. After this study, a definitive conclusion about the effect of the 3D-printed guidance system on the accuracy of chin repositioning can be drawn.

Appendices

A. Literature review

a. Abstract

Introduction: The goal of orthognathic surgery is to improve both function and facial appearance by correcting craniomaxillofacial deformities. The preparation of this surgery is conventionally based on two-dimensional (2D) analysis and planning. With the availability of three-dimensional (3D) imaging and software, computer-assisted surgery (CAS) has become a superior alternative to conventional planning. The planning of rotational and translational movements of the chin segment has become more detailed and extensive with the use of three-dimensional (3D) preoperative planning. For this reason, 3D-printed guidance systems have been developed to accurately transfer the detailed surgical plan to the intra-operative setting. The goal of this literature review is to put 3D-printed guidance systems for genioplasties in the context of orthognathic surgery and to give an overview of the current developments of the guidance systems.

Overview of CAS in orthognathic surgery: The workflows of 2D and 3D planning both consist of clinical facial analysis, cephalometry and surgical planning including soft tissue simulation. The 3D planning enables more elaborate diagnostics and surgical planning. For transferring the surgical plan to the operation room (OR), roughly five transfer categories are described: 1) freehand surgery, 2) traditional handmade acrylic splints, 3) 3D-printed splints, 4) navigation-assisted surgery and 5) splintless repositioning. During the postoperative accuracy assessment, it is evaluated whether the outcome of the surgery has improved. This process can be divided into two steps: 1) registration and 2) assessment.

Genioplasty: A small literature search in MEDLINE has identified ten studies. Even though each article described a different design, the systems are based on comparable principles. The transfer method is often a 3D-printed splint modified to guide the osteotomies as well. The guides are positioned using the bony contour and the occlusion as a reference. Commonly, two types of guides are used: one to perform the osteotomy and one to reposition the chin segment. Two clinical studies have compared the accuracy of chin repositioning with and without a guidance system, but neither were randomized trials.

Conclusion: For the postoperative accuracy assessment, the methodology of Baan et al. appears to be the method of choice. This method is based on voxel-based registration and the outcome parameters are indicative of changes in 3D. Published research on the 3D-printed guidance for genioplasties is limited. But a combination of 3D-printed cutting and repositioning guides, as described by Li et al, seems to be most promising. The lack of randomized controlled trials, combined with the predominance of Asian patients, supports the need for a prospective randomized controlled trial on 3D-printed guidance systems for genioplasties that includes Caucasian cases.

b. Introduction

The goal of orthognathic surgery is to improve both function and facial appearance by correcting craniomaxillofacial deformities. Patients with these deformities may experience functional problems, such as difficulties with chewing, biting and speaking [3]. In severe cases, skeletal deformities can lead to esthetic complaints in frontal view (asymmetry cases) or in sagittal view (strong concave or convex profile) [18]. Surgical techniques have been developed to adjust the skeleton to achieve functional and esthetic goals such as class I dental occlusion, facial balance and facial proportion [4]. Possible techniques to achieve the

optimal clinical result are repositioning of the maxilla or the mandible, by executing a Le Fort I osteotomy or bilateral sagittal split osteotomy (BSSO), respectively [4]. Depending on the patient's anatomy and pre-surgical orthodontic treatment, combinations of these techniques can be used to achieve an optimal result. Additional esthetic procedures such as genioplasty, rhinoplasty, zygoma-alterations are not uncommon. Repositioning of the jaws may produce big changes in the appearance of the patient. To guarantee a harmonious esthetic outcome, repositioning requires meticulous consultation, prediction, planning and psycho-social guidance before surgery.

Conventional planning is based on two-dimensional (2D) analysis and planning. With the availability of three-dimensional (3D) imaging and software, computer-assisted surgery (CAS) has become a superior alternative to conventional planning [8]. CAS is a broad concept that encompasses different techniques. The main characteristic is that part of the surgical process is supported by computer technology. CAS enables preoperative planning of the individual movements of the mandible, maxilla, and chin [9]. An advantage of CAS over conventional planning is that the surgeon can discuss the possibilities with the patient based on a 3D simulation, which improves the patient's appreciation compared to a 2D simulation [45]. Subsequently, the virtual plan can be transferred to the operation room (OR) with 3D-printed splints or navigation [46, 47]. Evaluation of the surgery's accuracy can also be performed with submillimeter accuracy using 3D imaging.

Together with the forehead, nose and zygomatic prominences, the chin plays a crucial part in the facial appearance [1]. The chin prominence is considered one of the fundamental pillars in completing facial harmony [48]. A chin osteotomy is executed for merely esthetic reasons, however it may functionally improve lip seal in open lip relations. It can be performed to correct a misalignment which is not corrected by a bimaxillary osteotomy, and can be used to increase facial harmony [1]. Before the availability of CAS, a surgical plan would consist of, for example, a 3 mm advancement of the chin segment. In a 3D planning however, the planned movement of the chin can consist of rotations and translations in all three dimensions, if this is necessary to correct the deformity. This potential increase in complexity of the surgical plan can pose a problem for the surgeon, who has to transfer the plan to the patient perioperatively. Therefore, distinct methods of transferring the increasingly complex surgical planning of genioplasty to the OR are being developed. One method to transfer the surgical planning of the chin is a recently introduced 3D-printed guidance system [12].

The goal of this literature review is to put 3D-printed guidance systems for genioplasties in the context of orthognathic surgery and to give an overview of the current developments of the guidance systems. Initially, an overview of CAS in orthognathic surgery is given. This consists of the two-dimensional (2D) and 3D planning methods, transfer of the planning to the OR and postoperative accuracy assessment. The focus of the second chapter lies on the genioplasty technique itself. A brief history is given and the basic concepts of osseous genioplasty are explained. Finally, the available literature on the transfer of the genioplasty surgical planning to the OR with its postoperative accuracy results is summarized.

c. Overview of computer-assisted surgery in orthognathic surgery

i. 2D and 3D planning methods

Planning of orthognathic surgery has rapidly evolved over the past decades. Conventionally, the planning is performed in 2D. This method is still widely used. The diagnostic phase consists of a clinical facial analysis and 2D cephalometric analysis of lateral radiographs (2D

x-rays) [13]. The results of the diagnostic phase are used to determine the repositioning during 2D surgical planning. Alternatively, planning can be performed in 3D. In this case, the diagnostic phase is based on a clinical facial analysis and a 3D cephalometric analysis of (CB)CT scans. Subsequently, the surgical planning is performed in 3D.

The first phase of both 2D and 3D planning is the clinical facial analysis. Facial analysis is performed for diagnostic reasons, for example, to determine the presence of a gummy smile or a sunday bite. The surgeon can judge whether the upper dentition, lower dentition, and chin are in the facial midline, whether teeth are aligned to the true vertical plane and whether the dental show is sufficient. It is also used to roughly determine the surgical plan, which may be altered in a later phase of the planning.

The next phase in 2D planning is cephalometric analysis. Cephalometric analysis was first introduced in the 1930s, by Hofrath [49] and Broadbent [50]. The literal meaning of cephalometry is ‘head measuring’. For this purpose, standardized lateral radiographs are traced [51]. In tracing, radiographic landmarks are identified, to construct lines, planes, and angles from which angular and linear values are computed. Figure 13 shows an example of such a tracing. The values are recorded and compared to a norm based on the patient’s ethnicity, age, and gender [52]. In this diagnostic phase, the surgeon can distinguish if, for example, a strong overbite is due to hyperplasia of the maxilla or hypoplasia of the mandible.

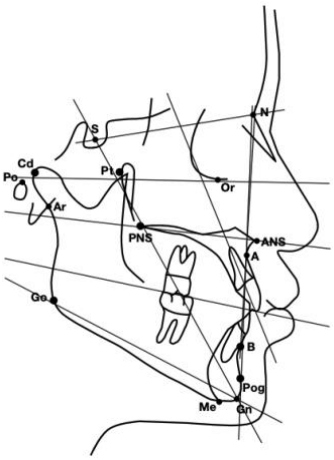


Figure 13 Cephalometric tracing. The identified landmarks are joined by lines to enable analysis of angular and linear relationships. From Hlongwa et al [51].

After the diagnostic phase is completed, the tracings are used for 2D surgical planning. The difference between the norm and the patient, together with the clinical analysis, are used to determine skeletal movements and a prediction of the surgical outcome can be made [53]. This prediction model is based on 2D surgical planning using either manual or digital tracings. Performing surgical planning based on manual tracings can be time-consuming, error-prone and subjective [51-53]. Because of its easily accessible application, the manual workflow is still widely used. However, digital methods have been developed to overcome the abovementioned drawbacks. Software packages can rapidly perform the analysis after indication of the landmarks. These programs require time, practice and precision to achieve a good accuracy of the predicted results [54]. Computer prediction programs such as Quick Ceph, Dolphin, DentoFacial Planner, etc., can achieve accurate precision outcomes of the soft tissue (less than 2 mm) [55].

The implementation of 3D planning was enabled by the introduction of Cone-Beam Computed Tomography (CBCT) in 1996 [56]. For 3D planning, the surgeon still needs to perform the abovementioned clinical facial analysis. Next, a 3D virtual model of the patient is required. The 3D virtual model is composed of a hard tissue (skeletal) model, segmented from a CBCT or Computed Tomography (CT) scan of the skeleton and a 3D scan of the dental arch [10]. Most software tools use global thresholding for the segmentation of the 3D models from the (CB)CT scans. In thresholding, a range of Hounsfield units (HU) or grey values, for CT or CBCT respectively, is selected to isolate the bone [14]. The resulting two models are merged into a 3D virtual model which resembles the patient as closely as possible [10]. A soft tissue model may be added from a segmentation of the (CB)CT scan. To complete this model, a stereophotograph can be overlaid in order to add integument which improves appreciation of the patient but is of no further value.

In analogy to the 2D planning workflow, the next phase in the 3D planning is cephalometric analysis. This requires identification of landmarks on the 3D virtual model. These landmarks can be indicated on different planes allowing computation of other distances or angles. Whereas 2D cephalometry only allows computation of angles in the sagittal view, 3D cephalometry enables computation of angles in all the three views (sagittal, frontal and transverse). In addition, it enables a comparison between left and right measures, such as the individual gonial angles [25]. Again, the computed values are compared to averages of healthy controls [25].

After the diagnostic phase, the surgical rotations and translations are determined in the 3D surgical planning. During this process, the esthetic effect of the surgical movement of the maxilla and/or mandible can be visualized and assessed, and the outcome of different surgical plans can be estimated both on bone and soft tissue level [45]. The level of detail of the planned translations is often submillimeter, with rotational accuracies of less than 1 degree.

The workflows of 2D and 3D planning roughly correspond. Both require a clinical facial analysis in combination with cephalometric analysis on imaging data. In both techniques, the effect on the hard and soft tissue are simulated in a way. Between the two, 3D planning enables more elaborate diagnostics and surgical planning. Due to the additional dimension and the ability to separately visualize and hide parts of the 3D model (for example the soft tissue, the mandible or maxilla), more extensive diagnostics can be performed. This allows for the detection of yaw deviations and other facial asymmetries which would not be detected by the 2D planning [10]. The additional dimension also increases the possibilities for the surgical planning. Whereas in 2D surgical planning, the surgeon primarily plans the anteroposterior and superior-inferior translation and the rotation in the sagittal plane (pitch), 3D surgical planning enables planning of the rotation in the coronal and axial plane (roll and yaw) and mediolateral translations.

One of the main issues in both 3D and 2D planning is the time required for the entire process. There is a lack of consensus on the change in time required for the planning. Some studies report that 3D virtual planning significantly decreases the required time compared to classic surgical planning [4, 57], whereas Pascal et al states that 3D virtual planning is time-consuming and can last up to 5 hours per case [45].

Another point of interest are the costs of the planning process. The least expensive option is 2D planning based on manual tracings, as this only requires the planner's time, not any software or expensive equipment. When 2D planning is based on digital tracings, the costs increase because of the required software. Performing the planning in 3D is considered the most expensive, since it requires computer equipment, software, and, sometimes, an engineer. According to literature, the entire setup can cost up to 80 000€ [45].

Lastly, the accessibility of both planning techniques differs. Whereas 3D planning is not available in every hospital, 2D planning is more easily accessible. This difference can be explained by the aforementioned high costs, but also by the time investment required for learning the technique and for the actual planning. In the Netherlands, surgeons have the possibility to outsource their surgery planning to other hospitals that have the right equipment and personnel available.

ii. Transfer of surgical planning to the operation room

Once the surgical plan is finished, the surgeon needs to execute the plan during the surgery. This requires a method to transfer the plan to the OR. There are a variety different methods described in literature, which can roughly be divided into five categories; 1) freehand surgery, 2) traditional handmade acrylic splints, 3) 3D-printed splints, 4) navigation-assisted surgery and 5) splintless repositioning [45].

1. Freehand surgery

In the freehand surgery technique, the surgeon estimates the osteotomized bone repositioning relying on his or her own experience and clinical evaluation in the OR. Some surgeons apply external reference points, such as the external reference nasal pin, to adequately control the vertical translation during Le Fort I surgery [58]. Advantageous of freehand surgery is the lack of additional costs and of time required for manufacturing a transfer tool [45]. However, there are a few important drawbacks. First, the execution of the plan depends solely on the surgeon's intraoperative assessment. The individual bone segments need to be repositioned with limited visibility. Next, the duration of the surgery may increase, since the surgeon needs to perform intraoperative measurements. Lastly, it is difficult to keep the bone segments in the correct position for fixation with osteosynthesis material [45].

2. Traditional handmade acrylic splints

This is probably the most frequently used transfer tool for orthognathic surgery [45]. Based on the determined repositioning in the surgical planning, the surgery is simulated using dental casts on articulators. This allows the manual fabrication of acrylic splints [59]. In single jaw cases, only one splint is fabricated, indicating the desired occlusion. In bimaxillary (double jaw) surgery, two splints are required: one intermediate and one final splint. The intermediate splint indicates the planned movement of the maxilla or mandible from the original position of its counterpart; the final splint indicates the movement of the second jaw and the desired occlusion (end-occlusion). Thus, the splint restores the relative position between the maxilla and mandible.

For a solitary BSSO, one splint is sufficient to completely transfer the surgical plan. However, during a Le Fort I surgery, the temporomandibular joint introduces a degree of freedom which impedes exact transfer. This is mainly reflected in the anteroposterior dimension [45]. In a bimaxillary surgery an additional potential source of error is introduced. In order to reposition

the maxillomandibular complex relative to the patient exactly according to the surgical plan, the reference frame of the dental casts in the articulators must match the reference frame of the patient. Since this matching is often not perfect, it results in discrepancies between achieved and planned rotations and translations [60]. The combination of both sources of errors can lead to an overall malpositioning of the jaws by up to 5 mm [61]. Also, since the splints are preoperatively fabricated, it is not possible to anticipate to unexpected events during the surgery. If the dentition changes or the surgeon decides on another final occlusion, a new splint is required. Beneficial of this transfer tool is that the tool itself does not require expensive equipment to create and that it increases the accuracy compared to freehand surgery [45].

3. 3D-printed splints

Multiple terms are used to describe this concept, such as Computer-Aided Design (CAD)/Computer-Aided Manufacturing (CAM), 3D-printed splints or patient-specific splints. These splints are designed, based on a virtual surgical plan, using 3D software, and are subsequently 3D-printed. Again, intermediate and final splints are fabricated. In contrast to the handmade splints, the 3D-printed splints are able to transfer the planned translations and rotations relative to the patient [45]. This is because the orientation of the 3D model is matched to the orientation of the patient. Yet, the anteroposterior error is still present due to the difference between pre- and postoperative condylar positioning. This is reflected in the cadaver study of Schouman et al., which reported an overall translational accuracy of the maxilla (which is dependent on the condylar positioning) of 1.55, 2.17 and 0.81 mm, for mediolateral, anteroposterior and superior-inferior respectively. The maximum overall rotational difference between planned and postoperative images was a pitch of 3.70 degrees of the maxilla [62]. A limitation of 3D-printed splints is that, as with handmade splints, there is no flexibility to change the surgical plan during surgery [45]. If this is necessary, the freehand surgery technique can be used as a fallback method. Another important drawback is the logistics of the fabrication of these splints. When in-house printing is unavailable, the 3D-printing needs to be outsourced. This requires extra planning time to ensure the splints are delivered well in advance of the surgery, to allow for fitting the splints on dental casts and for sterilization of the splints.

4. Navigation-assisted surgery

Navigation-assisted surgery, also known as surgical navigation, was introduced in the field of head and neck surgery more than 20 years ago. Over the years, it also gained acceptance for maxillofacial surgery [63]. The surgical navigation system requires four components: a localizer, a surgical probe, a CT scan dataset, and a visualization display. The most common type of navigation is based on an optical tracking system, in which light-reflecting markers are secured on the object to be tracked [63]. Different manufacturers such as Stryker, Medtronic, and Brainlab offer such systems. These systems require the registration of the CT data set to the actual patient. During this process, the system aligns the coordinate system of the patient to the CT scan. The systems can be differentiated based on their methods for registration, that is either point-based or surface-based matching. For example, the Stryker system requires identification of landmarks on the CT and the patient during surgery, which is a point-based technique [46]. The registration techniques will be explained in more detail in the subchapter ‘Postoperative accuracy assessment’.

After registration, the surgeon can see where the tracked object is located in the CT scan. For example, Mazzoni et al. tracked a surgical probe to evaluate the repositioning of the

maxilla and mandible. By including the 3D model of the desired position in the CT scan, the surgeon can intraoperatively check the repositioning by indicating reference points [46]. Figure 14 illustrates this process.

Beneficial of navigation is that no guides have to be manufactured. This reduces both the cost and preparation time of the surgery. Also, the anteroposterior accuracy is not decreased by the condylar repositioning. A total difference between the pre- and postoperative positioning of the maxilla of <0.61 mm is reported [46]. An important limitation of navigational surgery is the need for a dedicated, bulky and expensive system in the OR [45]. Also, the optical tracking system requires a free line-of-sight from the camera to the attached trackers. The accuracy of the navigation information is determined by the accuracy of the registration procedure; an inaccurate registration may affect accuracy and safety of the procedure. The total surgery time may increase with approximately 30-60 minutes due to the additional registration step and repeated checks during the surgery [46, 64]. In addition, stability of the bony segments before fixation remains a problem.

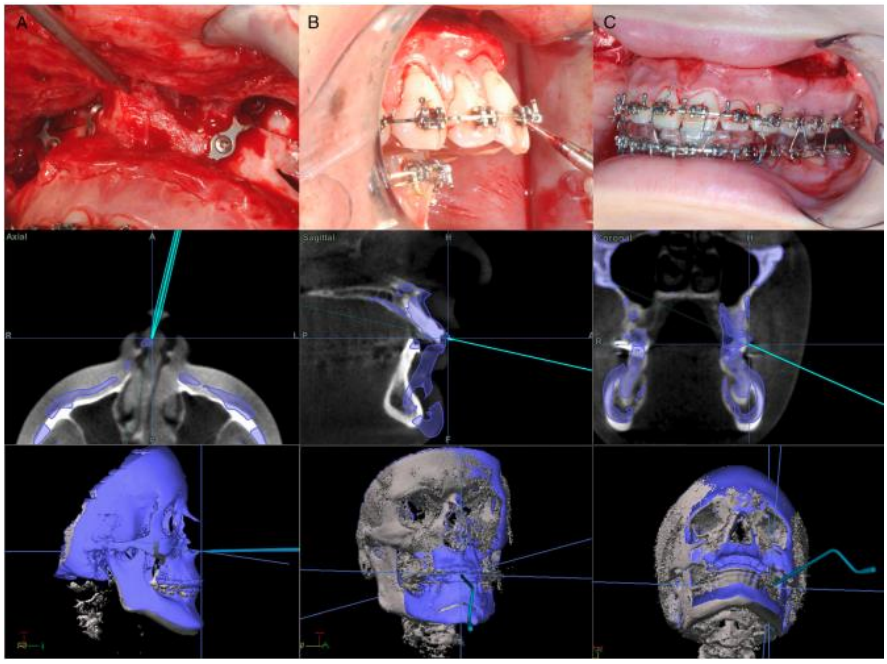


Figure 14 Intraoperative checks of three different reference points. A) anterior nasal spine, B) right central upper incisor, C) left first upper molar. From Mazzoni et al [46].

5. Splintless repositioning

A virtual surgical plan can be transferred to the OR without the application of splints. This requires predrilling/cutting guides in combination with custom-made fixation plates, as illustrated in Figure 15. The latter is also called a patient-specific implant (PSI). The guide is used to guide the osteotomy and predrill the holes required for fixation with the PSI. This technique could be used to move the mandible and maxilla separately, but most authors use it for the maxilla and reposition the mandible using a splint [45]. Advantageous of this method is that the repositioning is based on stable bone areas, eliminating the condylar repositioning error. Also, since this technique does not rely on the occlusion, it can be applied for edentulous patients. Next, the technique allows for a reduction in operating time, owing to the absence of intraoperative plate bending and intermaxillary fixation [60, 65]. Most importantly, this technique has a good accuracy, with a reported median deviation of the maxilla of the maxilla position between preoperative plan and surgical result of 0.39 mm [66].

The main limitations are the production costs of the PSIs. Since these are 3D-printed from titanium, the costs can be very high. The lack of flexibility during the surgery is again a limitation. The surgeon can only change the maxillary position by disregarding the customized plate, positioning the maxilla free-hand and manually bending a fixation plate [66]. At last, preformed plates seem to need wider surgical exposure to the upper jaw, contradicting needs for Minimal Invasive Orthognathic Procedures [67].

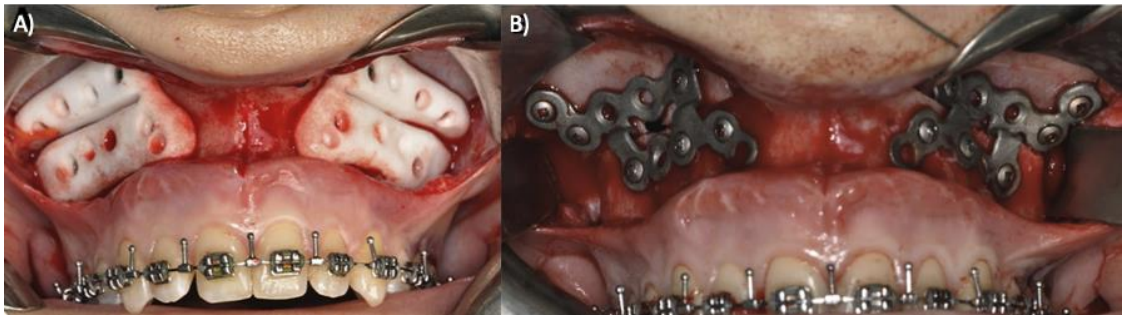


Figure 15 Surgery with A) predrilling/cutting guides and B) patient-specific implants, from Mazzoni et al [65].

iii. Postoperative accuracy assessment

Since much effort has been put into planning the surgery and transferring the plan to the OR, it is relevant to evaluate whether the outcome of the surgery has improved if more advanced techniques are utilized. This needs to be evaluated in the postoperative accuracy assessment. The workflow of the accuracy assessment can roughly be divided into two steps: 1) registration and 2) assessment [14]. For each step, different strategies are reported in the literature. These will be explained in the next paragraphs.

1. Registration techniques

To enable evaluation of the outcome, the preoperative planning and postoperative results need to be aligned. When describing alignment techniques, one often refers to the registration of two virtual models or image volumes. Different types of registration techniques are point-based registration [68], surface-based registration [9, 12, 22, 69-72] and voxel-based registration [15, 73]. For all registration techniques, the 'region of interest' for alignment is the part of hard tissue that is not affected by surgery, such as the anterior cranial base and frontal bone [14].

In point-based registration, reproducible points are indicated on both the fixed and moving image sets and their error distance is minimized. Traditionally, the result of orthognathic surgery was assessed in two dimensions. The pre- and postoperative lateral radiographs were registered by maximizing the overlap of the radiographic anatomical outlines [17]. With the availability of low dose CBCT imaging, registration techniques for 3D volumes have been developed. In 3D point-based registration, an algorithm fuses the fixed and moving models by minimizing the distance between the x, y, and z-coordinates of each point pair [68]. This process is prone to human error since the points are manually selected on each model [17]. Therefore, this method is rarely used.

In surface-based registration, a 3D surface model needs to be rendered from 3D volumes first. Each 3D model consists of a point cloud (vertices), interconnected by small triangles (faces). After a region of interest is defined, the Iterative Closest Point (ICP) approach is used to align the point clouds of the pre-operative and postoperative model. In this iterative process, the moving 3D point cloud is translated and rotated in each step, in order to minimize the point distances to the fixed 3D point cloud. Since CBCT is commonly used in maxillofacial surgery, the generation of a virtual 3D model relies on thresholding of grey values. It is reported that this can result in a

maximum difference of 1.2 mm between a CBCT surface model and regular CT surface model [14]. In addition, the 3D rendering step increases the processing time [17].

In voxel-based registration, the actual greyscale values are used to align the two 3D volumes. An algorithm, called mutual information (MI), searches the best match of the greyscale intensity between the voxels in the region of interest of both 3D volumes. MI is a measure of how well the intensity in the reference volume can be predicted, given the intensity in the moving volume [21]. Thus, the algorithm will maximize the MI by iteratively applying translations and rotations to the moving volume. Because all the information within the region of interest of the 3D volume is used, this technique requires more efficient computers and a longer processing time [17].

In conclusion, the study of Almukhtar et al. has compared the surface-based to the voxel-based technique in orthognathic surgery cases. It reported no statistically significant differences, but the voxel-based registration resulted in a lower variability [17]. This means the results were more consistent when voxel-based registration was applied.

2. Assessment

The second step entails the calculation of the actual difference between pre- and postoperative models. This may be computed in different ways. The first method is to compute the difference in linear and/or angular measurements compared to cephalometric landmarks [8, 22, 68]. Alternatively, one can compute the overall average difference between the two surfaces [22, 68, 70]. The last method is to compute the difference in the linear (in x, y, and z-direction) and angular (around x, y, and z-axis) movements of the 3D object [9, 12, 15, 71, 72].

For the latter technique, three points are manually indicated on the pre- and postoperative models, to form triangles that define the position and orientation of the models, as illustrated in Figure 16 [72]. The advantage of this technique is that it allows for evaluation of separate rotation and translation values. Baan et al. have developed a semi-automatic approach that combines voxel-based registration with the 3D position and orientation technique to evaluate the accuracy of orthognathic surgery [15]. Also, the human error, which can be introduced by identifying identical points on two models, is largely eliminated by selecting the three points only on the pre-operative model. Subsequently, the points are automatically identified in the postoperative model using voxel-based registration.

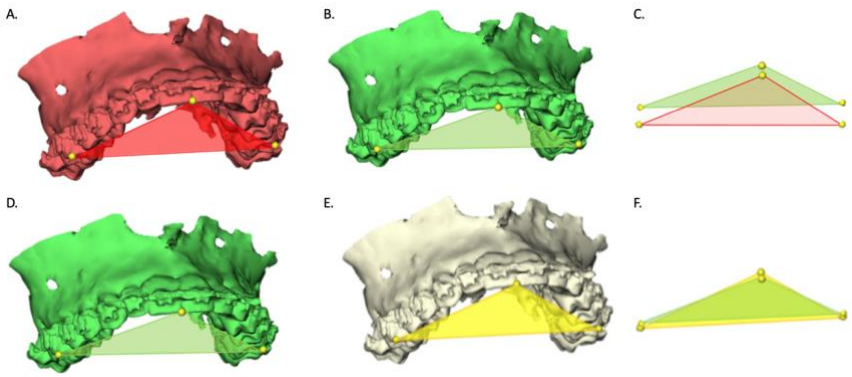


Figure 16 Visualization of the calculation with virtual triangles made up from three landmarks. Landmarks and virtual triangle were indicated on A) Preoperative virtually osteotomized maxilla and were translated to B) the 3D planned osteotomized maxilla by voxel-based registration. C) The difference between preoperative and planned position of the maxilla is displayed. Again, voxel-based registration was performed to transfer the landmarks and virtual triangle from D) the 3D planned maxilla to the E) postoperative maxilla. F) Difference between planned (green) and postoperative (yellow) position of the maxilla is displayed. From Baan et al [15].

In summary, combining the described techniques for the two steps can lead to multiple different methodologies to quantify the discrepancy between the planned and achieved position [14]. The goal of the systematic review of Gaber et al. was to reach an objective assessment protocol that could be universally used. Instead, it recommended the following; 1) perform voxel-based registration, 2) choose an outcome indicative of changes in 3D (translational/rotational in/around the different axes) and 3) compute inter- and intra-observer reliability to validate the results. These recommendations have led to the conclusion that OrthoGnathicAnalyser, developed by Baan et al [15], was the best method available at that moment

d. Genioplasty

i. A brief history

Bony genioplasty is traditionally used in correcting microgenia and asymmetry. Alternatives are inserting implants of fillers. Sparse comparative studies conclude that analysis of the results based on the entity of the chin's sagittal defect, the chin soft tissue thickness, the patient's age, and self-judgment allows for simplified treatment planning for sagittal chin deformities showing a greater predictability and a more stable long-term esthetic result regarding sliding genioplasty compared to alloplastic implant placement and fillers [74].

Early reports on genioplasty describe a submental skin incision to correct a chin that is placed too far back (retrogenia) by onlay techniques. Chin prominence was created by alloplastic implants and autologous grafts. Application of alloplastic materials enabled a natural shape, but complications such as displacement, infection, and erosion of the bony symphysis were common. The results of autologous grafts were quite good but temporary, as resorption was a main issue [75].

The first report describing an osseous genioplasty for a chin augmentation came from Hofer in 1942 [76]. The approach was extraoral and performed on a cadaver. In 1957, Trauner and Obwegeser described another technique to perform the osseous genioplasty. The intraoral approach was introduced, which allowed replacement of the chin while preserving its vascularity [77].

ii. Basic concepts of osseous genioplasty

There are multiple different surgical techniques to perform the osseous genioplasty. In this review, one mainstream technique is described [48]. The procedure starts with a local infiltration with anesthetic and vasoconstrictor. Next, a ‘step’ incision technique is performed. This entails a superficial mucosal incision from canine to canine, followed by a lower second incision through the mentalis muscle by pulling on the lip. This enables reattachment of the muscle to prevent a postoperative saggy chin. On both sides, the mental nerve is identified to prevent damage. Reference marks on the midline and possibly bilateral to the midline are placed to provide a reference for accurate repositioning. Different osteotomies (bone cuts) can be used to perform a genioplasty. They differ in shape of the osteotomy (straight line or a ‘box’ osteotomy) and its orientation. Each movement (i.e. lengthening or advancement) of the chin requires a different type of osteotomy. Some of the most common genioplasty techniques are discussed here:

- Advancement of the chin: A horizontal plane is required for the osteotomy. The shape of the mental area is determined by the height of the osteotomy (see Figure 17A). It is important to perform the osteotomy at least 5 mm below the root apex of the canine tooth, to secure vascularity and vitality of the dentition.

- Correction of vertical dimensions of the chin: a change in the angulation of the aforementioned horizontal plane will also affect the vertical dimension, as illustrated in Figure 17B. Another option is to perform an osteotomy of a segment of the genial bone segment or augmentation by down grafting, see Figure 17C-D.
- Correction of the transverse dimensions of the chin: a midline osteotomy is required for all cases. It depends on which aspect of the chin needs changing;
    - For the posterior aspect: a midline plate will be used as a pivot. When a widening is required, the defect can be filled with a bone graft (see Figure 18A). When a narrowing is required, a small triangular wedge is removed in the midline (see Figure 18C)
    - For the anterior aspect: When widening is required, the segments are separated with the predetermined by the midline plate. The defect can be filled with a bone graft (see Figure 18B). When a narrowing is required, a predetermined amount of bone is osteotomized. A midline plate is used to move and fixate the chin segments medially (see Figure 18D) [48].

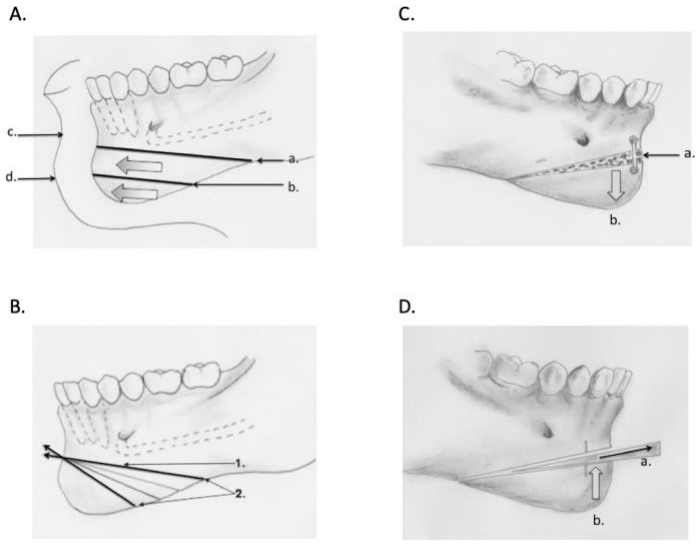


Figure 17 Correction of the anteroposterior and vertical dimensions of the chin. A) The height of the osteotomy determines the shape of the chin. B) When increasing the angle for a setback of the chin, the height of the chin will increase, whereas the advancement of the chin decreases the height of the chin. C) Increase of the height of the chin. D) Decrease of the height of the chin. From Ferretti et al [48].

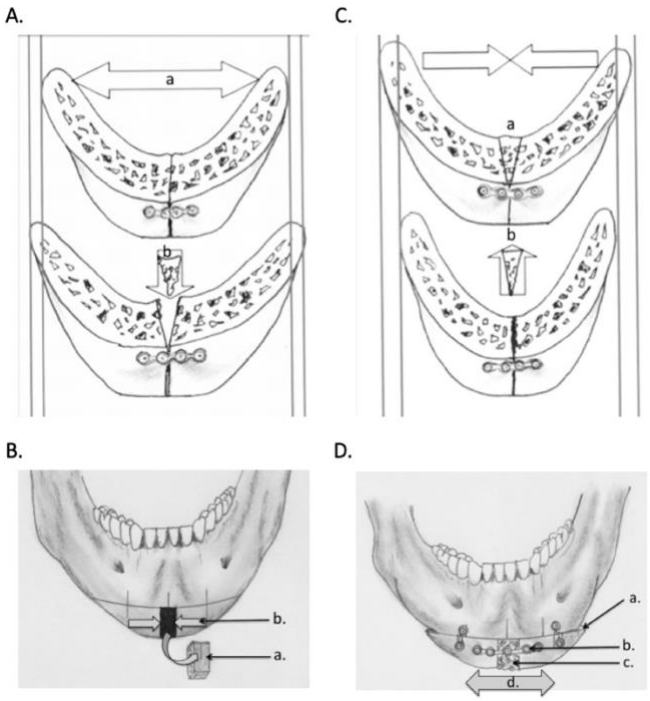


Figure 18 Correction of the transverse dimension of the chin. A) Increasing the width of the posterior aspect of the chin. Midline plate acts as a pivot. B) Narrowing anterior aspect of the chin. A predetermined amount of bone is removed, segments are moved medially. C) Decreasing the width of the posterior aspect of the chin. A small triangular segment is removed. Midline plate acts as a pivot. D) Widening anterior aspect of the chin. Segments are moved laterally and spaced apart with a midline plate. The defect can be filled with bone graft. From Ferretti et al [48].

iii. Transfer of surgical planning to the operation room

The application of 3D planning can potentially increase the complexity of the surgical planning. To help the surgeon in achieving even the most complex planning, several patient-specific guidance systems have been developed to aid in the execution of the genioplasty. A small literature search has been performed to identify studies concerning this subject. The following search strategy was used: *("Genioplasty"[Mesh] OR "Genioplasty"[title] OR "Genioglossus"[title]) AND ("Reproducibility of Results"[Mesh] OR "Accuracy"[tiab] OR "Accurate"[title]) AND ("Surgical Instruments"[Mesh] OR "Computer-Aided Design"[Mesh] OR "Genioplasty/instrumentation"[Mesh] OR "patient-specific"[title] OR "individual-specific"[title])*.

The process of selection and inclusion of articles is illustrated in the PRISMA-flow diagram [78] in Figure 19. The search in the MEDLINE database in March 2020 yielded 18 results. Reviewing references of other articles yielded four more articles. All 22 records were screened by title and abstract. Eight records were excluded because of the language (n = 1) or because no genioplasty was described (n = 7). The full texts of the remaining 14 articles, were retrieved and analyzed. Another four articles were excluded because the applied

guides were not 3D-printed (n = 2) or because a segmental genioplasty was performed (n = 2). In the end, a total of ten articles were included (Table 1).

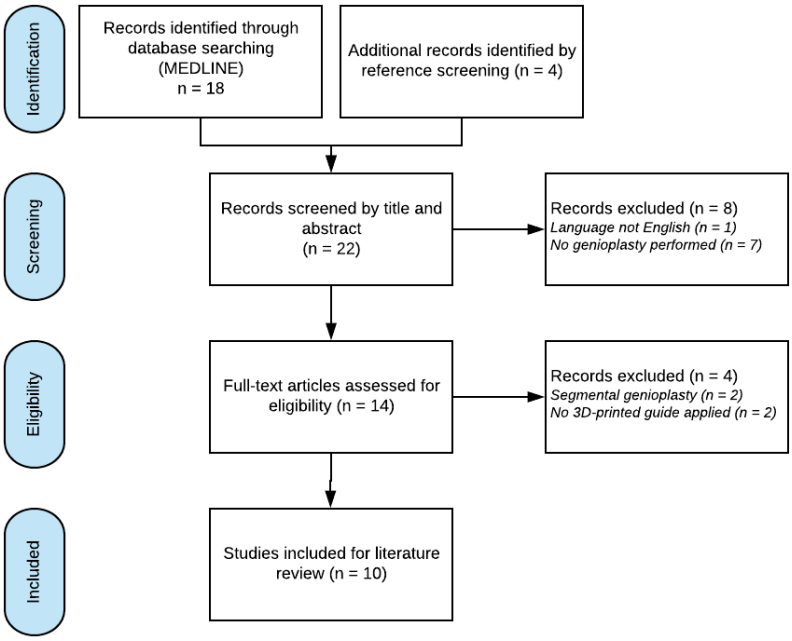


Figure 19 Flow diagram of the included articles

Table 15 Summary of the different designs of surgical guidance systems for genioplasties.

Study	Geography	Type of study	n (# with guides)	Type of guide	Stabilization of guide	Fixation of chin
Arcas et al. 2018	Spain	Case series	23 (23)	1 cutting guide 1 repositioning guide (from titanium)	Osteosynthesis screws Osteosynthesis screws	Repositioning guide, 4 predrilled holes used for fixation
Costa et al. 2018	Brazil	Technical note	1 (1)	1 drilling guide	Occlusion	Standard midline plate, 4 predrilled holes used for fixation
Hsu et al. 2013	USA	Prospective multicenter study	12 (8)	1 drilling guide 1 repositioning guide	Occlusion Occlusion + 2 osteosynthesis screws	Standard midline plate
Kang et al. 2014	Korea	In vitro prospective	30 (15)	1 cutting + repositioning guide	Osteosynthesis screws	NA
Li et al. 2017	China	Prospective control study	88 (44)	1 cutting guide 2 repositioning guides	Occlusion + osteosynthesis screws Osteosynthesis screws	Standard midline fixation plate
Lim et al. 2015	Korea	Case report	1 (1)	Uses design of Kang et al.; 1 cutting + repositioning guide	Osteosynthesis screws	Standard midline plate
Liu et al. 2017	USA	Case series	10 (10)	1 cutting guide 1 repositioning guide	Occlusion Occlusion + 2 osteosynthesis screws	Fixation plates
Polley et al. 2013	USA	Case report	1 (1)	1 drilling guide 1 positioning guide	Occlusion Occlusion + 2 osteosynthesis screws	Rigid fixation
Qiao et al. 2016	China	Case series	7 (7)	1 cutting guide 1 repositioning guide	Osseous relief of mandible Osseous relief of mandible	Lateral fixation with titanium mini plates.
Yamauchi et al. 2016	Japan	Technical note	1 (1)	1 cutting guide 1 repositioning guide	Occlusion Occlusion + 2 osteosynthesis screws	Positioning screws or titanium plate.

NA = not applicable

Even though each group uses a different design, the systems are based on comparable principles. The transfer method is often a 3D-printed splint (as mentioned in the ‘Orthognathic surgery - Transfer of surgical planning to the OR’ chapter), modified to guide

the osteotomies as well. The guides are positioned using the bony contour and the occlusion as a reference [9, 12, 79-82]. Osteosynthesis screws are often used for stabilization of the guide. In seven studies, two types of guides are used: one to perform the osteotomy and one to reposition the chin segment. As an example, Figure 20 shows the cutting and repositioning guides described by Li et al [12]. The other design types are described in the following paragraphs.

Costa et al. used one patient-specific guide [79]. They defined the final position of the chin in 3D modeling software (Dolphin Imaging, Chatsworth, USA) and fitted the STL of a standard osteosynthesis plate. This defined the correct positions of the holes of the fixation screws. After resetting the chin to the baseline position, the guide was designed to act as a drilling guide for the holes. The resulting guide is displayed in Figure 21A. Since a standardized plate is used, this concept will not work in more complex cases.

Kang et al. have designed a surgical guidance system that requires only one guide that both indicates the osteotomy line and the required repositioning, as can be seen in Figure 21B [83]. As the authors already reported, this system is only applicable to regular advancement or reduction genioplasties. This study describes the first results of mandibular models produced by rapid prototyping. Lim et al. have validated this exact system in actual patients for horizontal advancement of the chin [84].

The 3D printing material that was used to fabricate the surgical guide was not always mentioned. While the most common material reported was photosensitive resin [12, 82, 85], Arcas et al. manufactured the repositioning guide from titanium [86]. The benefit of titanium is that it allows the repositioning guide to act as a fixation plate as well. However, this does increase the overall costs of the procedure by 30% [86].

It is important to note that half of the identified studies took place in a (mostly) Asian patient population (see the second column in Table 1). It is known that the typical movement of the chin segment in Asian surgical cases (chin reduction) differs significantly from that in Caucasian cases (chin pronunciation) due to a difference in skeletal build. This justifies the need for a prospective study that includes Caucasian cases.

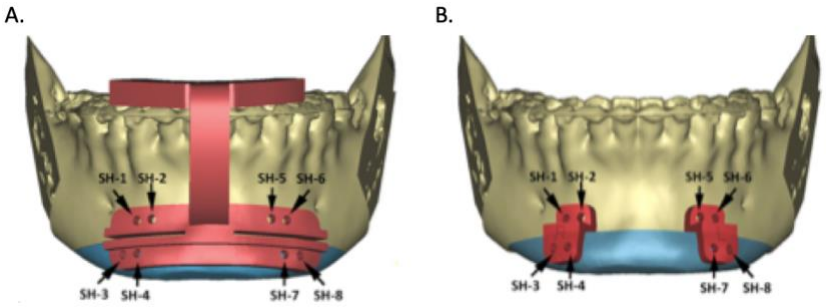


Figure 20 Genioplasty template system of Li et al [12]. A) upper portion of the osteotomy guide is designed like a dental splint. The lower portion indicates the osteotomy line. B) Pair of repositioning guides are installed using the Screw Holes (SH 1-8), this will automatically bring the chin segment into the final planned position.

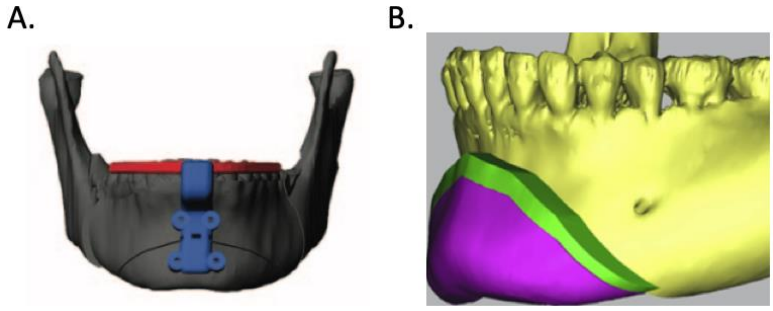


Figure 21 Genioplasty guidance systems consisting of one guide. A) The surgical guide of Costa et al [79]. B) The surgical guide of Kang et al [83].

iv. Postoperative accuracy results

Five studies reported quantitative accuracy outcomes. These results are summarized in Table 2. All studies used the surface-based registration technique (as described in the section 'Postoperative accuracy assessment' in the previous chapter). For the actual assessment, different techniques were applied. Costa et al. [79] reported the overall error from surface-to-surface, whereas Kang [83] and Qiao [85] computed differences based on cephalometric landmarks. The 3-points assessment technique was applied by Hsu [9] and Li et al [12]. No study has performed validation of their results by calculating inter- or intrarater correlations.

Only three studies have compared the accuracy of performing the genioplasty with and without surgical guides: two clinical studies and one in-vitro study. They have all reported a statistically significant improvement in the accuracy of the surgery [9, 12, 83].

The five studies reporting on quantitative accuracy outcomes are described in more detail below.

Table 16 Summary of the quantitative accuracy outcomes of genioplasties

Study	Registration technique	Accuracy assessment	Results					
			With guidance		Without guidance		Significance	
Costa et al 2018	Surface	Error from surface-to-surface: Overall RMSE [mm] Maximum error [mm]	 1.1 5.1		NA		NA	
Hsu et al 2013	Surface	3-points technique: RMSD of positional [mm] and rotational [degrees] differences of genial segment in x, y, z in relation to body of the mandible	 ML 0.6 AP 1.1 SI 0.5 Pitch 1.1 Roll 1.6 Yaw 1.6		 ML 1.4 AP 2.0 SI 1.5 Pitch 5.0 Roll 2.6 Yaw 2.9		For translation: F(3,20) = 9.29, p < 0.001 * For rotation: F(3,20) = 6.06, p = 0.004 *	
Kang et al 2014	Surface	From cephalometric landmarks: Absolute anterior transverse error (SD) Absolute anterior vertical errors (SD) Absolute posterior vertical error (SD)	*Advancement* 0.47 (0.35) 0.68 (0.49) 3.15 (2.08)	*Reduction* 0.57 (0.45) 0.27 (0.23) 3.14 (1.90)	*Advancement* 0.77 (0.45) 0.78 (0.53) 3.22 (1.92)	*Reduction* 0.62 (0.43) 0.58 (0.49) 3.84 (2.60)	*Advancement* p = 0.001 * p = 0.352 p = 0.865	*Reduction* p = 0.596 p < 0.001 * p = 0.152
Li et al 2017	Surface	3-points technique: RMSD of positional [mm] and rotational [degrees] differences for centroid of genial segment in x, y, z	 ML 0.62 AP 1.11 SI 0.84 Pitch 2.63 Roll 1.77 Yaw 1.51		 ML 1.45 AP 2.13 SI 2.63 Pitch 7.21 Roll 2.72 Yaw 2.92		"All 6 measurements were statistically significant"	
Qiao et al 2016	Surface	From cephalometric landmarks: Mean absolute position error /preoperative design movement [mm (%)]	 AP 0.15 (1.9) SI 0.06 (1.3)		NA		NA	

RMSD = root mean square deviation, RMSE = root mean square error, NA = not applicable, AP = anteroposterior, ML = mediolateral, SI = superior-inferior

* = statistically significant

1. In vitro study

Kang et al. has compared two genioplasty methods and verified their accuracies and efficacies on mandibular models which were produced by rapid prototyping [83]. In the experimental group, the genioplasty was executed with surgical guides, whereas the genioplasty was executed with manual measurements in the control group. The planning, which was equal for every case, was straightforward and consisted of a 5 mm advancement or 5 mm reduction of the chin. No rotations were planned (low complexity). Surface-based registration was applied to compare the postoperative to the planned models. The error at cephalometric landmarks, such as the symphysis, was calculated. For advancement genioplasty, the absolute anterior transverse error was 0.47 ± 0.35 (mean ± standard deviation) with the surgical guide, which was less than the error in the conventional method (0.77 ± 0.45; $p = 0.001$). For reduction genioplasty, the absolute anterior vertical error value was 0.27 ± 0.23 mm with the surgical guide versus 0.58 ± 0.49 mm with the conventional method ($p<0.001$). No quantification of rotations was performed. Extrapolation of these results to a clinical setting is difficult since the study was performed on a hard tissue cadaver model, with a free field of view and no soft tissue factors, as opposed to the clinical setting.

2. Clinical studies

In the study of Qiao et al. [85], a template system was designed for horizontal advancement genioplasty in 7 patients. No registration method was mentioned by the authors. Errors were calculated at cephalometric landmarks, such as the pogonion and menton. Only relative differences were described, no absolute measures were used. The reported relative errors were 1.9% and 1.3% for anteroposterior and superior-inferior translations, respectively. No comparison between the conventional method and the experimental method was made.

In the case study of Costa et al. [79], an iterative closest point algorithm is utilized for the surface-based registration. During this automatic alignment, a root mean square error of 1.1 mm from surface-to-surface was calculated. The resulting color map showed a maximum distance between the two surfaces of 5.1 mm.

Two clinical studies have compared the accuracy of chin repositioning with and without a guidance system [9, 12], but neither were randomized trials. Both used the 3-points technique to calculate the position and orientation of the segment. In the study of Hsu et al., the chin position and orientation were calculated in relation to the body of the mandible, to remove the confounding factor of the mandibular position. For the guidance group, the largest maximum root mean square deviation (RMSD) reported was 1.1 mm for the translation and 1.6 degrees, whereas 2.0 mm and 5.0 degrees were reported for the conventional group. The statistical analysis, performed with general linear models, resulted in statistically significant differences between the two groups for the translation ($p < 0.001$) and the rotation ($p = 0.004$) [9].

In the study of Li et al., the accuracies of conventional genioplasties and genioplasties executed with their genioplasty template system were compared [12]. A maximum RMSD of 1.1 in the anteroposterior direction and 2.6 degrees in pitch orientation for the experimental group were reported, versus 2.63 mm in superior-inferior direction and 7.21 degrees in pitch orientation for the control group. Their conclusion was that the template system provides greater accuracy.

e. Conclusion

The goal of orthognathic surgery is to improve both function and facial appearance by correcting dento-maxillofacial deformities. In preparation for the surgery, the individual

movements of the mandible, maxilla, and chin need to be planned. With the rise of 3D imaging and software, virtual surgical planning has become a valuable method for surgical planning. The application of 3D planning enables a more complex surgical planning when this is necessary to correct the deformity. Therefore, methods to transfer the potentially more complex planning have been developed. For the mandible and maxilla, five different categories of transfer methods for the surgical planning have been described in this review.

Concerning the postoperative accuracy assessment, the methodology of Baan et al. appears to be the method of choice. This is due to the fact that their method is based on voxel-based registration and the outcome variables are indicative of changes in 3D.

Published research on the 3D-printed guidance for genioplasties is limited. But a combination of 3D-printed cutting and repositioning guides, as described by Li et al, seems to be most promising. The lack of randomized controlled trials, combined with the predominance of Asian patients, supports the need for a prospective randomized controlled trial that includes Caucasian cases.

B. Ethical approval for the validation study

Amsterdam UMC
Universitair Medische Centra

Aan de heer prof. dr. A.G. Becking
Mondziekten, Kaak- en Aangezichtschirurgie
A1-115

Medisch Ethische Toetsingscommissie AMC
G4-214
telefoon: 020 56 67389

Amsterdam, 3 april 2020
uw kenmerk:
ons kenmerk: W20_127 # 20.159
betreft:
Uw project: **Validation of the OrthoGnathicAnalyser for assessing accuracy of genioplastic surgery**

Geachte heer Becking,

Uw brief d.d. 23 maart 2020 betreffende bovengenoemde studie is op 1 april jl. besproken in de vergadering van het dagelijks bestuur.

Het dagelijks bestuur is van oordeel dat bovengenoemde studie niet valt binnen de reikwijdte van de Wet medisch-wetenschappelijk onderzoek met mensen, aangezien er geen sprake is van wetenschappelijk onderzoek zoals bedoeld in artikel 1, eerste lid onder b van de WMO, daar de proefpersonen niet aan handelingen worden onderworpen, en/of aan de proefpersoon wordt geen bepaalde gedragswijze opgelegd, daar alleen retrospectief gegevens worden verzameld.

Een formele beoordeling door onze commissie is derhalve niet noodzakelijk.

De commissie attendeert u op de volgende punten:
De commissie heeft alleen de WMO-plichtigheid beoordeeld. Er heeft verder geen inhoudelijke toets van het onderzoek plaatsgevonden. U en uw afdeling zijn verantwoordelijk voor de correcte uitvoering van het onderzoek volgens de geldende wet- en regelgeving. Hierbij vragen wij uw aandacht voor de belangrijkste regelgeving:

- Voor prospectief onderzoek, waarbij gegevens van proefpersonen worden verzameld en verwerkt, is toestemming van de proefpersonen nodig.
- Voor retrospectief onderzoek, waarbij gegevens van proefpersonen gecodeerd worden verzameld en verwerkt is in beginsel toestemming van de proefpersonen nodig. In artikel 458 van de WGBO is vastgelegd onder welke omstandigheden van het vragen van toestemming kan worden afgezien. Bij retrospectief *anoniem* onderzoek is toestemming niet verplicht, hierbij zijn de gegevens nooit meer herleidbaar tot de proefpersonen. Dus ook niet via een code.
- Wanneer in een onderzoek gegevens worden verzameld van proefpersonen, dient hiermee correct te worden omgegaan zoals bepaald in de Gedragscode Gezondheidsonderzoek (Code Goed Gedrag), Algemene Verordening Gegevensbescherming en de Uitvoeringswet Algemene Verordening Gegevensbescherming en, indien het onderzoek van het AMC betreft, de regels die binnen het AMC zijn vastgesteld, zoals de SOP “Reuse of care data for the purpose of research van de CRU (http://kwadraet.amc.nl/iDocument/Viewers/Frameworks/ViewDocument.aspx?DocumentID=8efa33b4-9dfa-4575-a284-03b8fd8c7115&FromLogin=1)
- Wanneer in een onderzoek (lichaams)materiaal van proefpersonen wordt verzameld en verwerkt dient hiermee correct te worden omgegaan zoals bepaald in de Code Goed Gebruik. Indien er sprake is van een biobank in het AMC, dat wil zeggen dat het lichaamsmateriaal in het AMC wordt opgeslagen met het oog op toekomstig onderzoek, dient dit te worden voorgelegd aan de BiobankToetsingsCommissie (BTC) van het AMC.
- Onderzoek met anoniem materiaal vanuit de zorg is toegestaan, voorzover de patiënt van wie het materiaal afkomstig is hier geen bezwaar tegen heeft gemaakt (artikel 467 WGBO).
- Voorts dient u zich te houden aan de research code van het AMC en het VUmc.

AMC en VUmc werken samen in Amsterdam UMC

AMC
Meibergdreef 9
1105 AZ Amsterdam
Postbus 22660
1100 DD Amsterdam
T +31(0)20 566 9111
www.amc.nl

Meer informatie over bovengenoemde regelgeving kunt u vinden op internet, waaronder onze intranetpagina. Deze opsomming betreft de belangrijkste regelgeving, maar is niet uitputtend. Mogelijk is nog andere wet- en regelgeving van toepassing op uw onderzoek.

Indien u twijfelt of door amendering of het toevoegen van addenda het onderzoek nog steeds buiten de reikwijdte van de WMO blijft kunt u dit aan de commissie ter beoordeling voorleggen.

Met vriendelijke groet,
namens de Medisch Ethische Toetsingscommissie AMC,

Drs. O. Harlaar
ambtelijk secretaris

Zo lang de beperkende maatregelen als gevolg van het coronavirus gelden zullen de brieven inzake beoordeling WMO-plichtigheid van de METC niet worden voorzien van een natte handtekening. Deze brieven worden digitaal verstuurd. Indien u na het intrekken van de maatregelen alsnog een ondertekende brief nodig heeft, verneemt de METC dit graag.

Bijlage: verklaring in het Engels

c.c. per email: j.f.sabelis@amsterdamumc.nl; r.schreurs@amsterdamumc.nl; t.c.vanriet@amsterdamumc.nl; thomas.maal@radboudumc.nl

AMC en VUmc werken samen in Amsterdam UMC

AMC
Meibergdreef 9
1105 AZ Amsterdam

Postbus 22660
1100 DD Amsterdam

T +31(0)20 566 9111
www.amc.nl

References

1. Abadi M, Pour OB. Genioplasty. Facial Plast Surg. 2015;31(5):513-22.
2. Ward JL, Garri JI, Wolfe SA. The osseous genioplasty. Clin Plast Surg. 2007;34(3):485-500.
3. Oh J, Han JJ, Ryu SY, Oh HK, Kook MS, Jung S, et al. Clinical and Cephalometric Analysis of Facial Soft Tissue. J Craniofac Surg. 2017;28(5):e431-e8.
4. Naran S, Steinbacher DM, Taylor JA. Current Concepts in Orthognathic Surgery. Plast Reconstr Surg. 2018;141(6):925e-36e.
5. Seah TE, Bellis H, Ilankovan V. Orthognathic patients with nasal deformities: case for simultaneous orthognathic surgery and rhinoplasty. Br J Oral Maxillofac Surg. 2012;50(1):55-9.
6. Liddle MJ, Baker SR, Smith KG, Thompson AR. Psychosocial Outcomes in Orthognathic Surgery: A Review of the Literature. Cleft Palate Craniofac J. 2015;52(4):458-70.
7. Posnick JC, Choi E, Chang RP. Osseous genioplasty in conjunction with bimaxillary orthognathic surgery: a review of 262 consecutive cases. Int J Oral Maxillofac Surg. 2016;45(7):904-13.
8. De Riu G, Virdis PI, Meloni SM, Lumbau A, Vaira LA. Accuracy of computer-assisted orthognathic surgery. J Craniomaxillofac Surg. 2018;46(2):293-8.
9. Hsu SS, Gateno J, Bell RB, Hirsch DL, Markiewicz MR, Teichgraeber JF, et al. Accuracy of a computer-aided surgical simulation protocol for orthognathic surgery: a prospective multicenter study. J Oral Maxillofac Surg. 2013;71(1):128-42.
10. Stokbro K, Aagaard E, Torkov P, Bell RB, Thygesen T. Virtual planning in orthognathic surgery. Int J Oral Maxillofac Surg. 2014;43(8):957-65.
11. DICOM standard, C.7.6.2.1.1. Image Position and Image Orientation: National Electrical Manufacturers Association (NEMA); [Available from: http://dicom.nema.org/medical/dicom/current/output/chtml/part03/sect_C.7.6.2.html#sect_C.7.6.2.1.1.
12. Li B, Wei H, Zeng F, Li J, Xia JJ, Wang X. Application of A Novel Three-dimensional Printing Genioplasty Template System and Its Clinical Validation: A Control Study. Sci Rep. 2017;7(1):5431.
13. Haas OL, Jr., Becker OE, de Oliveira RB. Computer-aided planning in orthognathic surgery-systematic review. Int J Oral Maxillofac Surg. 2014.
14. Gaber RM, Shaheen E, Falter B, Araya S, Politis C, Swennen GRJ, 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(11):2430-40.
15. Baan F, Liebregts J, Xi T, Schreurs R, de Koning M, Berge S, et al. A New 3D Tool for Assessing the Accuracy of Bimaxillary Surgery: The OrthoGnathicAnalyser. PLoS One. 2016;11(2):e0149625.
16. Streiner DL, Norman GR. "Precision" and "accuracy": two terms that are neither. J Clin Epidemiol. 2006;59(4):327-30.
17. Almukhtar A, Ju X, Khambay B, McDonald J, Ayoub A. 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(4):e93402.
18. Jung HD, Kim SY, Park HS, Jung YS. Orthognathic surgery and temporomandibular joint symptoms. Maxillofac Plast Reconstr Surg. 2015;37(1):14.
19. Gateno J, Xia J, Teichgraeber JF, Rosen A, Hultgren B, Vadnais T. The precision of computer-generated surgical splints. J Oral Maxillofac Surg. 2003;61(7):814-7.
20. Swennen GR, Mollemans W, Schutyser F. Three-dimensional treatment planning of orthognathic surgery in the era of virtual imaging. J Oral Maxillofac Surg. 2009;67(10):2080-92.
21. Yoo TS. Insight Into Images: Principles and Practice for Segmentation, Registration, and Image Analysis: A K Peters; 2004.
22. Zavattero E, Romano M, Gerbino G, Rossi DS, Gianni AB, Ramieri G, et al. Evaluation of the Accuracy of Virtual Planning in Orthognathic Surgery: A Morphometric Study. J Craniofac Surg. 2019;30(4):1214-20.
23. Stokbro K, Liebregts J, Baan F, Bell RB, Maal T, Thygesen T, et al. Does Mandible-First Sequencing Increase Maxillary Surgical Accuracy in Bimaxillary Procedures? J Oral Maxillofac Surg. 2019;77(9):1882-93.
24. Liebregts J, Baan F, van Lierop P, de Koning M, Bergé S, Maal T, et al. One-year postoperative skeletal stability of 3D planned bimaxillary osteotomies: maxilla-first versus mandible-first surgery. Sci Rep. 2019;9(1):3000.
25. Swennen GRJ, Schutyser, F., Hausamen, J.E. Three-Dimensional Cephalometry: A Color Atlas and Manual: Springer-Verlag Berlin Heidelberg; 2006.
26. Maes F, Collignon A, Vandermeulen D, Marchal G, Suetens P. Multimodality image registration by maximization of mutual information. IEEE Trans Med Imaging. 1997;16(2):187-98.
27. Nada RM, Maal TJ, Breuning KH, Bergé SJ, Mostafa YA, Kuijpers-Jagtman AM. Accuracy and reproducibility of voxel based superimposition of cone beam computed tomography models on the anterior cranial base and the zygomatic arches. PLoS One. 2011;6(2):e16520.
28. Shaheen E, Shujaat S, Saeed T, Jacobs R, Politis C. Three-dimensional planning accuracy and follow-up protocol in orthognathic surgery: a validation study. Int J Oral Maxillofac Surg. 2019;48(1):71-6.
29. 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(6):1316-26.

30. Titiz I, Laubinger M, Keller T, Hertrich K, Hirschfelder U. Repeatability and reproducibility of landmarks--a three-dimensional computed tomography study. Eur J Orthod. 2012;34(3):276-86.
31. Swennen GR, Schutyser F, Barth EL, De Groeve P, De Mey A. A new method of 3-D cephalometry Part I: the anatomic Cartesian 3-D reference system. J Craniofac Surg. 2006;17(2):314-25.
32. Katkar RA, Kummet C, Dawson D, Moreno Uribe L, Allareddy V, Finkelstein M, et al. Comparison of observer reliability of three-dimensional cephalometric landmark identification on subject images from Galileos and i-CAT cone beam CT. Dentomaxillofac Radiol. 2013;42(9):20130059.
33. Torosdagli N, Liberton DK, Verma P, Sincan M, Lee JS, Bagci U. Deep Geodesic Learning for Segmentation and Anatomical Landmarking. IEEE Trans Med Imaging. 2019;38(4):919-31.
34. de Jong MA, Gül A, de Gijt JP, Koudstaal MJ, Kayser M, Wolvius EB, et al. Automated human skull landmarking with 2D Gabor wavelets. Phys Med Biol. 2018;63(10):105011.
35. Hariton E, Locascio JJ. Randomised controlled trials - the gold standard for effectiveness research: Study design: randomised controlled trials. Bjog. 2018;125(13):1716.
36. Onderzoek CCM. Your research: Is it subject to the WMO or not? : Centrale Commissie Mensgebonden Onderzoek; 2020 [Available from: https://english.ccmo.nl/investigators/legal-framework-for-medical-scientific-research/your-research-is-it-subject-to-the-wmo-or-not.
37. Medical Devices: Guidance Document - Classification of medical devices 2.4/1 Rev. 9, (2010).
38. Leone FC, Nelson, L., Nottingham, R. . The Folded Normal Distribution. Techometrics. 1961;3(4):543-50.
39. Broglio K. Randomization in Clinical Trials: Permuted Blocks and Stratification. Jama. 2018;319(21):2223-4.
40. Berger VW, Grant, W.C. Randomization Procedures. Wiley StatsRef: Statistics Reference Online; 2014.
41. Berger VW, Antsigyna, O. A review of randomization methods in clinical trials. Open Access J 2015(Future Science Ltd).
42. Kahan BC, Rehal S, Cro S. Risk of selection bias in randomised trials. Trials. 2015;16:405-.
43. Cunningham SJ, Garratt AM, Hunt NP. Development of a condition-specific quality of life measure for patients with dentofacial deformity: II. Validity and responsiveness testing. Community Dent Oral Epidemiol. 2002;30(2):81-90.
44. Chen K, Sun X, Wang L, Ni S, Wu Z, Gao Y, et al. A Modified Cosmetic Genioplasty Can Affect Airway Space Positively in Skeletal Class II Patients: Studying Alterations of Hyoid Bone Position and Posterior Airway Space. Aesthetic Plast Surg. 2020.
45. Pascal E, Majoufre C, Bondaz M, Courtemanche A, Berger M, Bouletreau P. Current status of surgical planning and transfer methods in orthognathic surgery. J Stomatol Oral Maxillofac Surg. 2018;119(3):245-8.
46. Mazzoni S, Badiali G, Lancellotti L, Babbi L, Bianchi A, Marchetti C. Simulation-guided navigation: a new approach to improve intraoperative three-dimensional reproducibility during orthognathic surgery. J Craniofac Surg. 2010;21(6):1698-705.
47. Li B, Shen S, Jiang W, Li J, Jiang T, Xia JJ, 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(10):1298-305.
48. Ferretti C, Reyneke JP. Genioplasty. Atlas Oral Maxillofac Surg Clin North Am. 2016;24(1):79-85.
49. Hofrath ODH. Die bedeutung der röntgenfern-und abstand-saufnahme für die diagnostik der kieferanomalien. Fortschritte der Orthodontik in Theorie und Praxis. 1931;1(2):232-58.
50. Broadbent BH. A new X-ray technique and its application to orthodontia. The Angle Orthodontist. 1931;1(2):45-66.
51. Hlongwa P. Cephalometric analysis: manual tracing of a lateral cephalogram. SADJ. 2019;Vol. 74:318-22.
52. Duterloo HS. A reflection on radiographic cephalometry: the evaluation of sagittal discrepancy. J Orthod. 2014;41(3):208-17.
53. Marianetti TM, Gasparini G, Midulla G, Grippaudo C, Deli R, Cervelli D, et al. Numbers of Beauty: An Innovative Aesthetic Analysis for Orthognathic Surgery Treatment Planning. Biomed Res Int. 2016;2016:6156919.
54. Kusnoto B. Two-dimensional cephalometry and computerized orthognathic surgical treatment planning. Clin Plast Surg. 2007;34(3):417-26.
55. Kaipatur NR, Flores-Mir C. Accuracy of computer programs in predicting orthognathic surgery soft tissue response. J Oral Maxillofac Surg. 2009;67(4):751-9.
56. Pittayapat P, Limchaichana-Bolstad N, Willems G, Jacobs R. Three-dimensional cephalometric analysis in orthodontics: a systematic review. Orthod Craniofac Res. 2014;17(2):69-91.
57. Steinhuber T, Brunold S, Gartner C, Offermanns V, Ulmer H, Ploder O. Is Virtual Surgical Planning in Orthognathic Surgery Faster Than Conventional Planning? A Time and Workflow Analysis of an Office-Based Workflow for Single- and Double-Jaw Surgery. J Oral Maxillofac Surg. 2018;76(2):397-407.
58. Ruckman P, 3rd, Schlieve T, Borba AM, Miloro M. External Reference Nasal Pin for Orthognathic Maxillary Positioning: What Is the Proper Method of Placement? J Oral Maxillofac Surg. 2016;74(2):399.e1-9.
59. Hammoudeh JA, Howell LK, Boutros S, Scott MA, Urata MM. Current Status of Surgical Planning for Orthognathic Surgery: Traditional Methods versus 3D Surgical Planning. Plast Reconstr Surg Glob Open. 2015;3(2):e307.

60. Van den Bempt M, Liebregts J, Maal T, Berge S, Xi T. Toward a higher accuracy in orthognathic surgery by using intraoperative computer navigation, 3D surgical guides, and/or customized osteosynthesis plates: A systematic review. J Craniomaxillofac Surg. 2018;46(12):2108-19.
61. Ellis E, 3rd. Accuracy of model surgery: evaluation of an old technique and introduction of a new one. J Oral Maxillofac Surg. 1990;48(11):1161-7.
62. Schouman T, Rouch P, Imholz B, Fasel J, Courvoisier D, Scolozzi P. Accuracy evaluation of CAD/CAM generated splints in orthognathic surgery: a cadaveric study. Head Face Med. 2015;11:24.
63. Azarmehr I, Stokbro K, Bell RB, Thygesen T. Surgical Navigation: A Systematic Review of Indications, Treatments, and Outcomes in Oral and Maxillofacial Surgery. J Oral Maxillofac Surg. 2017;75(9):1987-2005.
64. Zinser MJ, Sailer HF, Ritter L, Braumann B, Maegele M, Zoller JE. A paradigm shift in orthognathic surgery? A comparison of navigation, computer-aided designed/computer-aided manufactured splints, and "classic" intermaxillary splints to surgical transfer of virtual orthognathic planning. J Oral Maxillofac Surg. 2013;71(12):2151.e1-21.
65. Mazzoni S, Bianchi A, Schiariti G, Badiali G, Marchetti C. Computer-aided design and computer-aided manufacturing cutting guides and customized titanium plates are useful in upper maxilla waferless repositioning. J Oral Maxillofac Surg. 2015;73(4):701-7.
66. Heufelder M, Wilde F, Pietzka S, Mascha F, Winter K, Schramm A, et al. Clinical accuracy of waferless maxillary positioning using customized surgical guides and patient specific osteosynthesis in bimaxillary orthognathic surgery. J Craniomaxillofac Surg. 2017;45(9):1578-85.
67. Brunso J, Franco M, Constantinescu T, Barbier L, Santamaria JA, Alvarez J. Custom-Machined Miniplates and Bone-Supported Guides for Orthognathic Surgery: A New Surgical Procedure. J Oral Maxillofac Surg. 2016;74(5):1061.e1-.e12.
68. Zinser MJ, Mischkowski RA, Sailer HF, Zoller JE. Computer-assisted orthognathic surgery: feasibility study using multiple CAD/CAM surgical splints. Oral Surg Oral Med Oral Pathol Oral Radiol. 2012;113(5):673-87.
69. Hernandez-Alfaro F, Guijarro-Martinez R. New protocol for three-dimensional surgical planning and CAD/CAM splint generation in orthognathic surgery: an in vitro and in vivo study. Int J Oral Maxillofac Surg. 2013;42(12):1547-56.
70. Lin HH, Chang HW, Lo LJ. Development of customized positioning guides using computer-aided design and manufacturing technology for orthognathic surgery. Int J Comput Assist Radiol Surg. 2015;10(12):2021-33.
71. Stokbro K, Aagaard E, Torkov P, Bell RB, Thygesen T. Surgical accuracy of three-dimensional virtual planning: a pilot study of bimaxillary orthognathic procedures including maxillary segmentation. Int J Oral Maxillofac Surg. 2016;45(1):8-18.
72. Xia JJ, Gateno J, Teichgraeber JF, Christensen AM, Lasky RE, Lemoine JJ, et al. Accuracy of the computer-aided surgical simulation (CASS) system in the treatment of patients with complex craniomaxillofacial deformity: A pilot study. J Oral Maxillofac Surg. 2007;65(2):248-54.
73. Tucker S, Cevidanes LH, Styner M, Kim H, Reyes M, Proffit W, et al. Comparison of actual surgical outcomes and 3-dimensional surgical simulations. J Oral Maxillofac Surg. 2010;68(10):2412-21.
74. Bertossi D, Galzignato PF, Albanese M, Botti C, Botti G, Nocini PF. Chin Microgenia: A Clinical Comparative Study. Aesthetic Plast Surg. 2015;39(5):651-8.
75. Obwegeser HL. Orthognathic surgery and a tale of how three procedures came to be: a letter to the next generations of surgeons. Clin Plast Surg. 2007;34(3):331-55.
76. Hofer O. Operation del Prognathie und Microgenie. Dtsch Zahn Mund Kleferhlkd. 1942;9:121-32.
77. Trauner R, Obwegeser H. The surgical correction of mandibular prognathism and retrognathia with consideration of genioplasty. I. Surgical procedures to correct mandibular prognathism and reshaping of the chin. Oral Surg Oral Med Oral Pathol. 1957;10(7):677-89; contd.
78. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 2009;6(7):e1000097.
79. Costa PJC, Demetrio MS, Nogueira P, Rodrigues LR, Junior PDR. A New Proposal for Three-Dimensional Positioning of the Chin Using a Single Computer-Aided Design/Computer-Aided Manufacturing Surgical Guide. J Craniofac Surg. 2018;29(7):1963-4.
80. Liu SY, Huon LK, Zaghi S, Riley R, Torre C. An Accurate Method of Designing and Performing Individual-Specific Genioglossus Advancement. Otolaryngol Head Neck Surg. 2017;156(1):194-7.
81. Polley JW, Figueroa AA. Orthognathic positioning system: intraoperative system to transfer virtual surgical plan to operating field during orthognathic surgery. J Oral Maxillofac Surg. 2013;71(5):911-20.
82. Yamauchi K, Yamaguchi Y, Katoh H, Takahashi T. Tooth-bone CAD/CAM surgical guide for genioplasty. Br J Oral Maxillofac Surg. 2016;54(10):1134-5.
83. Kang SH, Lee JW, Lim SH, Kim YH, Kim MK. Validation of mandibular genioplasty using a stereolithographic surgical guide: in vitro comparison with a manual measurement method based on preoperative surgical simulation. J Oral Maxillofac Surg. 2014;72(10):2032-42.
84. Lim SH, Kim MK, Kang SH. Genioplasty using a simple CAD/CAM (computer-aided design and computer-aided manufacturing) surgical guide. Maxillofac Plast Reconstr Surg. 2015;37(1):44.
85. Qiao J, Fu X, Gui L, Girod S, Lee GK, Niu F, et al. Computer Image-Guided Template for Horizontal Advancement Genioplasty. J Craniofac Surg. 2016;27(8):2004-8.

86. Arcas A, Vendrell G, Cuesta F, Bermejo L. Advantages of performing mentoplasties with customized guides and plates generated with 3D planning and printing. Results from a series of 23 cases. J Craniomaxillofac Surg. 2018;46(12):2088-95.