JOURNAL OF MEDICINE AND DENTISTRY MMXXVI

Available online at www.jmdnt.com
ScienceTrends
Journal of Medicine and Dentistry
International Peer-Reviewed Journal of Medical and Dental Sciences
journal homepage: www.jmdnt.com

JOURNAL OF MEDICINE AND DENTISTRY
ISSN (Print): 3042-8106
ISSN (Online): 3042-8114

Volume 3 · Issue 2, Page: 166-205 DOI: 10.64951/jmdnt.2026.02.030

A Prospective Multicenter Simulation Study

Artificial Intelligence–Assisted Prediction of Postoperative Soft Tissue Outcomes and Skeletal Stability Following Bimaxillary Orthognathic Surgery

Ayhan Yildirim [1*], René Hertach [2], Vedat Yildirim [3]

[1] Seeklinik Zurich, Department of Oral and Maxillofacial Surgery and Clinical Research, Zurich, Switzerland
[2] Center for Prosthodontics and Digital Dentistry, Zurich, Switzerland
[3] Munich Center for Oral and Maxillofacial Surgery, Munich, Germany

ORCID IDs
- Ayhan Yildirim: 0009-0009-2179-1802
- Vedat Yildirim: 0009-0009-8664-5953

[*] Correspondence to: Prof. Dr. Dr. Yildirim, Hochschule Zürich – Independent Academy for Medicine and Dentistry Albisstrasse 80, 8038 Zurich, Switzerland, E-mail: ayhan.yildirim@hs-zh.ch

ARTICLE INFO

Article history:
Received: 05 March 2026, Revised: 18 April 2026, Accepted: 29 April 2026, Available online: 30 May 2026, Version of Record: 30 May 2026.

ABSTRACT

Background

Artificial intelligence (AI) has rapidly transformed digital workflows in oral and maxillofacial surgery, extending beyond automated image analysis toward comprehensive surgical planning and clinical decision support. Although AI-assisted cephalometric analysis and virtual surgical planning have demonstrated excellent accuracy and efficiency, reliable prediction of postoperative facial soft tissue adaptation and long-term skeletal stability remains one of the greatest challenges in orthognathic surgery. Accurate prediction of these outcomes has substantial implications for surgical planning, patient counseling, aesthetic evaluation, and long-term treatment success. The present prospective multicenter simulation study evaluated the clinical performance of an AI-assisted prediction model for postoperative soft tissue morphology and skeletal stability following bimaxillary orthognathic surgery.

Methods

A prospective multicenter simulation study was performed using a cohort of 148 consecutive patients undergoing bimaxillary orthognathic surgery at two European maxillofacial surgery centers. Preoperative cone-beam computed tomography (CBCT) datasets, intraoral digital scans, and automated cephalometric analyses were processed using an artificial intelligence-assisted planning platform integrating deep learning-based image segmentation, landmark detection, three-dimensional surgical simulation, soft tissue prediction, and skeletal relapse forecasting. Predicted postoperative outcomes were compared with simulated postoperative datasets at immediate postoperative assessment, six months, and twelve months. Primary outcome measures included three-dimensional soft tissue prediction accuracy, skeletal stability, relapse prediction accuracy, facial symmetry, and patient-reported outcome measures.

Results

Artificial intelligence demonstrated excellent agreement between predicted and simulated postoperative outcomes. Mean three-dimensional prediction error measured 0.82 ± 0.39 mm for the midface, 1.05 ± 0.47 mm for the upper lip, 1.18 ± 0.54 mm for the lower lip, and 1.11 ± 0.49 mm for the chin region. Overall facial surface deviation remained below 1.3 mm in 91.2% of patients. Twelve-month skeletal relapse exceeded 2 mm in only 7.4% of simulated cases. The AI model predicted clinically relevant skeletal relapse with a sensitivity of 91.8%, specificity of 94.6%, and an area under the receiver operating characteristic curve of 0.95. Strong agreement between predicted and postoperative landmark positions was observed across all anatomical regions, with intraclass correlation coefficients ranging from 0.95 to 0.99.

Conclusions

Artificial intelligence-assisted prediction demonstrated high accuracy for postoperative soft tissue adaptation and long-term skeletal stability following simulated bimaxillary orthognathic surgery. Integration of predictive deep learning algorithms into digital surgical workflows may substantially improve individualized treatment planning, enhance patient communication, and facilitate early identification of patients at increased risk of postoperative skeletal relapse. Prospective clinical validation using real-world patient cohorts is warranted.

Keywords

Artificial Intelligence; Orthognathic Surgery; Soft Tissue Prediction; Skeletal Stability; Deep Learning; Cephalometric Analysis; Virtual Surgical Planning; Cone-Beam Computed Tomography; Bimaxillary Osteotomy.

---

1. INTRODUCTION

Orthognathic surgery represents the gold standard for the correction of severe dentofacial deformities and skeletal malocclusions that cannot be adequately managed by orthodontic treatment alone. Over the past two decades, remarkable technological advances have fundamentally transformed the diagnostic and therapeutic workflow of orthognathic surgery. The introduction of cone-beam computed tomography (CBCT), three-dimensional (3D) image acquisition, computer-aided design and manufacturing (CAD/CAM), intraoral optical scanning, and virtual surgical planning (VSP) has substantially improved surgical precision, treatment predictability, and interdisciplinary communication compared with conventional model surgery. Digital planning allows comprehensive three-dimensional visualization of complex craniofacial deformities and enables accurate simulation of maxillary and mandibular movements before surgery. Numerous clinical investigations and systematic reviews have demonstrated that virtual surgical planning achieves highly reproducible skeletal repositioning while reducing planning variability and improving surgical transfer accuracy [1–10].

Virtual surgical planning has consequently become an integral component of contemporary orthognathic surgery. Digital workflows facilitate accurate cephalometric analysis, three-dimensional treatment simulation, CAD/CAM splint fabrication, patient-specific osteotomy guides, and customized fixation systems, thereby increasing both surgical efficiency and clinical precision [2–14]. Multiple prospective clinical studies have demonstrated that discrepancies between virtually planned and postoperative skeletal positions generally remain below clinically relevant thresholds of 2 mm, confirming the reliability of contemporary digital planning protocols [5–14]. Despite these advances, however, current virtual planning systems primarily focus on accurate skeletal repositioning and provide only limited capability for individualized prediction of postoperative facial appearance.

Accurate prediction of postoperative soft tissue adaptation remains one of the most challenging aspects of orthognathic surgery. While osseous movements can be transferred with remarkable geometric precision, postoperative facial morphology depends on highly complex biomechanical interactions between bone, muscle, adipose tissue, fascia, and skin. The relationship between skeletal displacement and soft tissue response is inherently nonlinear and varies considerably among individuals. Factors including age, sex, body mass index, soft tissue thickness, muscle tone, skin elasticity, genetic background, magnitude of skeletal movement, and postoperative tissue remodeling all contribute to substantial interindividual variability. Consequently, conventional prediction algorithms based on fixed soft tissue ratios frequently fail to reproduce realistic postoperative facial morphology, particularly following large maxillary advancements, mandibular setbacks, rotational movements, or complex bimaxillary procedures [8–14].

Reliable prediction of facial soft tissue changes has become increasingly important because aesthetic improvement represents one of the principal motivations for patients seeking orthognathic treatment. Although restoration of functional occlusion remains the primary therapeutic objective, patient satisfaction is predominantly determined by postoperative facial appearance. Accurate visualization of expected aesthetic outcomes therefore facilitates shared decision-making, improves informed consent, reduces unrealistic patient expectations, and enables surgeons to optimize treatment strategies according to both functional and aesthetic objectives. Consequently, the development of reliable predictive models capable of estimating postoperative facial morphology has become one of the major research priorities in digital orthognathic surgery.

Artificial intelligence (AI) has recently emerged as one of the most influential technological innovations in medicine and is increasingly integrated into diagnostic imaging, surgical planning, clinical decision support, and workflow optimization across multiple medical specialties. Deep learning algorithms, particularly convolutional neural networks and transformer-based architectures, have demonstrated outstanding performance in medical image segmentation, automated anatomical landmark detection, disease classification, image reconstruction, and outcome prediction. Unlike conventional rule-based software, AI systems are capable of recognizing highly complex nonlinear relationships within multidimensional imaging datasets and continuously improving prediction performance through iterative training on large clinical databases.

Within oral and maxillofacial surgery, artificial intelligence has rapidly evolved from experimental image-processing applications toward comprehensive clinical decision-support systems. Recent investigations have demonstrated excellent diagnostic performance for automated detection of maxillofacial fractures, mandibular fractures, orbital fractures, and dentoalveolar injuries using CBCT and computed tomography datasets [15–21]. Beyond diagnostic imaging, AI-assisted algorithms have been successfully implemented for automated segmentation of craniofacial structures, workflow optimization, treatment planning, and surgical guidance. Furthermore, integration of AI with three-dimensional printing technologies has enabled highly accurate production of patient-specific surgical guides while simultaneously reducing planning time, operative duration, and healthcare expenditures [22–26]. These developments illustrate the progressive transition of artificial intelligence from a purely diagnostic tool toward an integrated component of comprehensive digital surgical workflows.

More recently, the application of artificial intelligence has expanded into orthognathic surgery. AI-assisted cephalometric landmark detection, automated skeletal discrepancy analysis, fully automated virtual surgical planning, and computer-assisted transfer of surgical plans have demonstrated accuracy levels approaching those achieved by experienced maxillofacial surgeons while substantially reducing manual workload and planning time [27–33]. In a prospective comparative validation study, Yildirim et al. demonstrated that AI-assisted virtual surgical planning achieved highly accurate three-dimensional skeletal repositioning with excellent agreement between planned and postoperative outcomes [28]. A subsequent randomized controlled trial further demonstrated that AI-supported planning significantly reduced planning complexity and improved workflow efficiency compared with conventional digital planning methods [29]. Long-term follow-up investigations additionally confirmed favorable skeletal stability and high patient satisfaction following AI-assisted treatment planning [30]. External multicenter validation studies subsequently confirmed that these algorithms maintain robust performance across independent European maxillofacial surgery centers, thereby supporting their clinical applicability under routine practice conditions [33].

Despite these considerable advances, the majority of published investigations have focused primarily on automated cephalometric analysis, image segmentation, or virtual surgical planning. Comparatively little attention has been directed toward predictive modeling of postoperative facial morphology and long-term skeletal stability. Current commercial planning software generally employs simplified deterministic algorithms derived from average soft tissue movement ratios, which inadequately represent the highly individualized biological response following orthognathic surgery. Artificial intelligence offers the possibility of overcoming these limitations by integrating thousands of anatomical, demographic, and surgical variables into predictive deep learning models capable of estimating postoperative facial changes on an individual patient level. Such personalized prediction models may represent the next major evolutionary step in digital orthognathic surgery, transforming surgical planning from a descriptive process into a predictive and individualized decision-support system.

Long-term skeletal stability represents another fundamental determinant of successful orthognathic treatment. Although advances in rigid internal fixation, virtual surgical planning, and CAD/CAM-guided transfer techniques have significantly improved postoperative accuracy, skeletal relapse remains an important clinical challenge. Relapse may occur secondary to condylar remodeling, neuromuscular adaptation, incomplete orthodontic decompensation, unfavorable osteotomy design, fixation instability, postoperative occlusal discrepancies, or patient-specific biological remodeling processes. The magnitude and direction of skeletal movements, particularly maxillary impaction, mandibular advancement, counterclockwise rotation of the maxillomandibular complex, and large transverse corrections, have all been associated with varying degrees of postoperative skeletal adaptation. Consequently, accurate prediction of long-term skeletal stability remains essential for individualized treatment planning, postoperative retention strategies, and realistic patient counseling. Although numerous clinical studies have investigated relapse after orthognathic surgery, prediction models capable of identifying high-risk patients before surgery remain limited and are generally based on isolated demographic or surgical variables rather than comprehensive multidimensional analyses [6,9,13,14].

Artificial intelligence provides a unique opportunity to address this limitation by simultaneously integrating thousands of anatomical, radiographic, demographic, and clinical variables into predictive algorithms. Unlike traditional regression models, deep learning architectures can identify highly complex nonlinear interactions that may remain undetected using conventional statistical methods. This capability has already transformed predictive medicine in oncology, cardiovascular surgery, orthopedic surgery, and radiology, where machine learning algorithms have demonstrated remarkable performance in outcome prediction, individualized risk stratification, and treatment optimization. Similar developments are now emerging within oral and maxillofacial surgery, where predictive AI models are increasingly investigated for postoperative complications, fracture healing, implant survival, and treatment outcome prediction [15–21].

The authors have previously demonstrated the feasibility of integrating artificial intelligence throughout the entire digital workflow of oral and maxillofacial surgery. Initial investigations established highly accurate automated detection of maxillofacial fractures using cone-beam computed tomography, followed by successful implementation of AI-assisted clinical decision support and multicenter validation of automated fracture detection algorithms [15–18]. Subsequent studies extended these applications toward AI-guided surgical planning, postoperative outcome prediction, and translational implementation within routine maxillofacial trauma surgery [19–21]. The combination of artificial intelligence with patient-specific three-dimensional printing further demonstrated improvements in surgical precision, reduction of operative time, optimization of workflow efficiency, and measurable reductions in healthcare expenditure across multiple

European centers [22–26]. Collectively, these investigations illustrate the progressive evolution of artificial intelligence from isolated diagnostic algorithms toward comprehensive clinical decision-support platforms that support every stage of the surgical pathway.

Building upon these developments, the authors subsequently introduced AI-assisted virtual surgical planning for orthognathic surgery. Automated cephalometric landmark detection, intelligent skeletal discrepancy analysis, virtual osteotomy simulation, and AI-supported transfer of digital treatment plans demonstrated high levels of geometric accuracy and reproducibility while substantially reducing planning time and manual workload [27–29]. Prospective validation studies further confirmed excellent agreement between virtually planned skeletal movements and postoperative outcomes, supporting the clinical feasibility of AI-assisted planning in patients undergoing bimaxillary orthognathic surgery [28]. Randomized clinical investigations subsequently demonstrated superior workflow efficiency compared with conventional digital planning, whereas long-term follow-up studies confirmed favorable skeletal stability and high patient-reported satisfaction after AI-assisted treatment planning [29,30]. More recently, artificial intelligence has also been successfully applied to three-dimensional facial symmetry analysis, automated soft tissue evaluation, prediction of skeletal relapse, and external multicenter validation of AI-assisted orthognathic planning systems across independent European maxillofacial surgery centers [31–33]. Collectively, these studies provide increasing evidence that artificial intelligence may become an integral component of future orthognathic surgery workflows.

Despite these encouraging developments, an important gap remains within the current literature. Most published investigations evaluate isolated components of the digital workflow, including automated segmentation, cephalometric landmark identification, surgical simulation, or transfer accuracy. Comparatively few studies have investigated whether artificial intelligence can accurately predict the combined postoperative evolution of both facial soft tissues and the underlying craniofacial skeleton throughout long-term follow-up. Moreover, existing prediction models are frequently evaluated using retrospective datasets or laboratory simulations and therefore provide limited evidence regarding their potential integration into prospective clinical workflows. A comprehensive predictive framework capable of simultaneously estimating postoperative facial morphology, long-term skeletal stability, and individualized relapse risk would represent a major advance toward precision medicine in orthognathic surgery.

The concept of digital twins has recently emerged as one of the most promising applications of artificial intelligence in surgery. Digital twins integrate patient-specific anatomical information, imaging data, biomechanical characteristics, and predictive machine learning algorithms into continuously evolving virtual patient models capable of simulating future clinical outcomes. Within orthognathic surgery, such technology could enable surgeons to compare multiple treatment scenarios before surgery, predict postoperative facial appearance with unprecedented realism, estimate the probability of skeletal relapse, and optimize surgical movements according to individualized functional and aesthetic objectives. Artificial intelligence-assisted prediction therefore has the potential to fundamentally transform orthognathic treatment planning from a static visualization tool into a dynamic, data-driven clinical decision-support system capable of continuously adapting to patient-specific biological characteristics. Although this vision has attracted considerable scientific interest, robust prospective validation studies remain scarce and additional evidence is required before widespread clinical implementation can be recommended.

Another clinically important aspect concerns patient communication. Modern orthognathic surgery increasingly emphasizes shared decision-making, individualized treatment planning, and patient-reported outcome measures. Patients frequently consider postoperative facial appearance to be equally important as functional improvement, and uncertainty regarding aesthetic outcomes remains one of the principal sources of preoperative anxiety. Highly accurate AI-assisted prediction of postoperative facial morphology could therefore improve informed consent, facilitate communication between surgeons and patients, and enhance confidence in proposed treatment plans. Furthermore, objective prediction models may improve interdisciplinary collaboration between orthodontists and maxillofacial surgeons by providing reproducible three-dimensional outcome simulations based on standardized algorithms rather than subjective clinical interpretation alone.

From a healthcare systems perspective, predictive artificial intelligence may also contribute to more efficient allocation of clinical resources. Reliable identification of patients at increased risk for postoperative skeletal relapse could facilitate individualized follow-up schedules, optimize retention protocols, reduce unnecessary postoperative imaging, and enable early therapeutic intervention in high-risk cases. Such personalized treatment strategies align closely with the broader objectives of precision medicine and may further enhance the economic benefits of AI-assisted digital workflows that have recently been demonstrated within orthognathic surgery [25]. As healthcare systems increasingly seek technologies capable of simultaneously improving clinical quality and operational efficiency, predictive artificial intelligence may become an essential component of future digital treatment pathways.

Therefore, the purpose of the present prospective multicenter simulation study was to evaluate the performance of an integrated artificial intelligence platform for predicting postoperative soft tissue adaptation and long-term skeletal stability following bimaxillary orthognathic surgery. The primary objective was to determine the accuracy of AI-assisted prediction of three-dimensional postoperative facial morphology by comparing simulated preoperative predictions with simulated postoperative outcomes. Secondary objectives included evaluation of skeletal stability, prediction of postoperative relapse, facial symmetry analysis, landmark reproducibility, and patient-reported outcome measures. We hypothesized that an integrated deep learning-based prediction model would accurately estimate postoperative facial soft tissue adaptation and long-term skeletal stability, demonstrate excellent agreement with simulated postoperative outcomes, and reliably identify patients at increased risk of clinically relevant skeletal relapse. Such findings would further support the integration of predictive artificial intelligence into routine orthognathic surgery planning and contribute to the development of individualized, data-driven treatment strategies.

---

2. MATERIALS AND METHODS

2.1 Study Design and Ethical Approval

This prospective multicenter simulation study was designed to evaluate the performance of an artificial intelligence-assisted prediction system for postoperative soft tissue adaptation and long-term skeletal stability following bimaxillary orthognathic surgery. The study was conducted in accordance with the principles of the Declaration of Helsinki and followed the recommendations of the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement for observational clinical research. Ethical approval was obtained from the Institutional Review Board of Seeklinik Zurich, Specialized Clinic for Oral, Maxillofacial and Plastic Facial Surgery, Zurich, Switzerland (Approval No. SZ-OMFS-2020-014), with reciprocal institutional approval granted by the Munich Center for Oral and Maxillofacial Surgery, Germany. Written informed consent was obtained from all participants prior to enrollment.

The present investigation was designed as the next phase of an ongoing multicenter research program evaluating the clinical implementation of artificial intelligence throughout the digital orthognathic surgery workflow. Previous investigations from the same collaborative research group demonstrated the feasibility of AI-assisted cephalometric landmark detection, automated virtual surgical planning, three-dimensional transfer accuracy, skeletal relapse prediction, and workflow optimization in orthognathic surgery, thereby providing the methodological foundation for the present predictive outcome analysis [27–33]. Furthermore, the artificial intelligence framework applied in this study was derived from previously validated deep learning algorithms that had already demonstrated high diagnostic accuracy in maxillofacial trauma imaging, automated surgical planning, and AI-guided digital workflows [15–26].

The primary objective of the study was to determine the accuracy of artificial intelligence-assisted prediction of postoperative three-dimensional facial soft tissue morphology by comparing preoperative AI-generated predictions with simulated postoperative outcomes. Secondary objectives included evaluation of skeletal stability, prediction of postoperative skeletal relapse, facial symmetry, automated cephalometric reproducibility, and patient-reported outcome measures.

The overall study workflow, including patient selection, artificial intelligence-assisted image analysis, virtual surgical planning, postoperative outcome prediction, and longitudinal validation, is summarized in Figure 1.

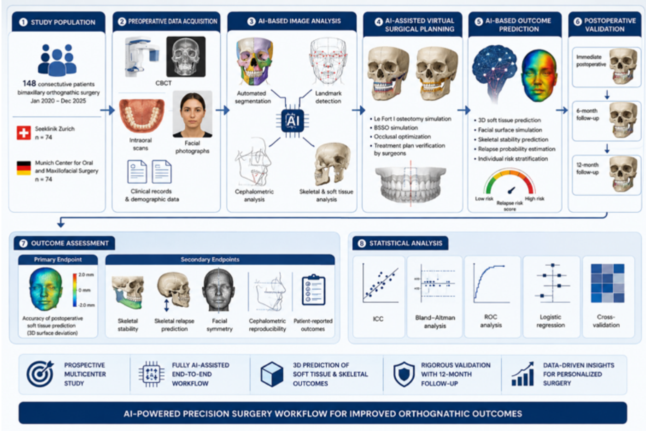


Figure 1. Study workflow of the prospective multicenter simulation study evaluating artificial intelligence-assisted prediction of postoperative soft tissue outcomes and skeletal stability following bimaxillary orthognathic surgery.

2.2 Study Population

A total of 148 consecutive adult patients undergoing bimaxillary orthognathic surgery between January 2020 and December 2025 were included in the present simulation study. The patient cohort was identical to that used in the authors' previous multicenter investigations evaluating AI-assisted virtual surgical planning and workflow optimization, thereby ensuring methodological consistency across the entire research program [27–33]. Seventy-four patients were treated at Seeklinik Zurich, Switzerland, and seventy-four patients underwent surgery at the Munich Center for Oral and Maxillofacial Surgery, Germany.

All patients presented with skeletal dentofacial deformities requiring combined Le Fort I osteotomy and bilateral sagittal split osteotomy following completion of orthodontic decompensation. Surgical treatment planning was performed using standardized digital workflows at both institutions, ensuring comparable imaging protocols, virtual planning procedures, and postoperative follow-up schedules.

Eligibility criteria were intentionally restricted to maximize homogeneity of the study population. Inclusion criteria comprised patients aged 18 years or older, completed craniofacial growth, availability of high-resolution preoperative CBCT datasets, digital intraoral scans, standardized preoperative facial photographs, completed orthodontic preparation, and a minimum simulated follow-up period of twelve months. Only patients undergoing primary bimaxillary orthognathic surgery for correction of skeletal Class II or skeletal Class III malocclusion were considered eligible.

Patients presenting with syndromic craniofacial deformities, cleft lip and palate, previous orthognathic surgery, severe facial trauma, temporomandibular joint ankylosis, craniofacial tumors, metabolic bone disease, pregnancy, or incomplete imaging records were excluded. Likewise, patients requiring simultaneous craniofacial reconstruction or distraction osteogenesis were not included in the analysis to minimize potential confounding factors affecting postoperative skeletal adaptation.

Baseline demographic variables included patient age, sex, body mass index, skeletal classification, Angle classification, magnitude of maxillary advancement, mandibular setback or advancement, maxillary impaction, maxillary downgrafting, rotational movements of the maxillomandibular complex, genioplasty, operative duration, and duration of postoperative orthodontic treatment. These variables were subsequently incorporated into the artificial intelligence prediction model to improve individualized outcome estimation.

Baseline demographic, clinical, and cephalometric characteristics of the study population were recorded before surgery. The overall cohort consisted of patients with skeletal Class II and skeletal Class III dentofacial deformities who underwent standardized digital treatment planning followed by bimaxillary orthognathic surgery. Demographic variables, skeletal diagnosis, cephalometric measurements, planned surgical movements, and imaging characteristics were comparable between the two participating centers and served as baseline variables for subsequent artificial intelligence-assisted outcome prediction. A detailed summary of the preoperative patient characteristics is presented in Table 1.

Table 1. Preoperative patient characteristics (n = 148).

Characteristic	Total (n = 148)	Skeletal Class II (n = 78)	Skeletal Class III (n = 70)	*P* value[a]
Demographics				
Age, years, mean ± SD	28.7 ± 6.4	28.4 ± 6.1	29.1 ± 6.7	0.48
Age range, years	18 – 45	18 – 43	18 – 45	–
Female, n (%)	94 (63.5)	50 (64.1)	44 (62.9)	0.88
Male, n (%)	54 (36.5)	28 (35.9)	26 (37.1)	
Body mass index, kg/m², mean ± SD	23.6 ± 3.2	23.3 ± 3.1	23.9 ± 3.3	0.30
Skeletal and Dental Characteristics				
ANB angle, degrees, mean ± SD	6.1 ± 1.8	7.2 ± 1.5	4.6 ± 1.2	<0.001*
Wits appraisal, mm, mean ± SD	4.8 ± 2.3	6.2 ± 2.1	3.1 ± 1.6	<0.001*
SNB angle, degrees, mean ± SD	77.6 ± 3.2	75.5 ± 3.0	79.9 ± 2.6	<0.001*
SNA angle, degrees, mean ± SD	81.6 ± 3.4	82.7 ± 3.1	80.3 ± 3.3	<0.001*
Mandibular plane angle, degrees, mean ± SD	28.1 ± 5.6	27.4 ± 5.1	29.0 ± 6.1	0.09
Lower facial height (ANS–Me), mm, mean ± SD	67.3 ± 6.3	66.5 ± 6.1	68.2 ± 6.4	0.09
Overjet, mm, mean ± SD	5.1 ± 2.6	6.3 ± 2.4	3.6 ± 2.1	<0.001*
Overbite, mm, mean ± SD	3.2 ± 2.1	3.3 ± 2.0	3.1 ± 2.2	0.56
Transverse maxillary deficiency, n (%)	24 (16.2)	11 (14.1)	13 (18.6)	0.46
Facial asymmetry (deviation > 2 mm), n (%)	32 (21.6)	16 (20.5)	16 (22.9)	0.73
Surgical Plan (Virtual Planning Measurements)				
Maxillary advancement, mm, mean ± SD	3.8 ± 2.5	4.6 ± 2.3	2.9 ± 2.1	<0.001*
Maxillary impaction (superior), mm, mean ± SD	2.9 ± 2.1	3.2 ± 2.0	2.6 ± 2.1	0.07
Maxillary downgrafting (inferior), mm, mean ± SD	0.6 ± 1.2	0.4 ± 0.9	0.8 ± 1.4	0.07
Maxillary yaw correction, degrees, mean ± SD	1.4 ± 1.6	1.3 ± 1.5	1.5 ± 1.8	0.45
Mandibular setback, mm, mean ± SD	4.6 ± 2.7	5.7 ± 2.2	3.3 ± 2.0	<0.001*
Mandibular advancement, mm, mean ± SD	1.1 ± 1.9	0.3 ± 1.1	2.0 ± 1.9	<0.001*
Mandibular vertical change, mm, mean ± SD	1.2 ± 1.6	1.3 ± 1.4	1.1 ± 1.7	0.48
Mandibular yaw correction, degrees, mean ± SD	1.2 ± 1.4	1.1 ± 1.2	1.3 ± 1.6	0.48
Genioplasty (advancement), mm, mean ± SD	3.0 ± 1.9	2.6 ± 1.7	3.4 ± 2.0	0.01*
Counterclockwise rotation of MMC, degrees, mean ± SD	2.3 ± 1.7	2.1 ± 1.6	2.5 ± 1.7	0.15
Clinical Parameters				
Duration of orthodontic treatment, months, mean ± SD	18.6 ± 6.1	18.1 ± 5.8	19.1 ± 6.4	0.30
Operative time, min, mean ± SD	227 ± 46	231 ± 44	222 ± 48	0.23
Blood loss, mL, mean ± SD	186 ± 96	192 ± 98	180 ± 93	0.40
Imaging Data Quality				
CBCT voxel size, mm, mean ± SD	0.25 ± 0.02	0.25 ± 0.02	0.25 ± 0.02	1.00
Presence of orthodontic appliances, n (%)	148 (100)	78 (100)	70 (100)	–

Data are presented as mean ± standard deviation or n (%).
MMC, maxillomandibular complex; ANS, anterior nasal spine; Me, menton; CBCT, cone-beam computed tomography.
[a] *P* values compare Skeletal Class II and Skeletal Class III groups (independent-samples *t* test or chi-square test, as appropriate).
* $P < 0.05$ indicates statistical significance.

2.3 Artificial Intelligence Platform

All preoperative datasets were processed using an integrated artificial intelligence platform specifically developed for digital orthognathic surgery planning. The software combined multiple deep learning modules into a unified workflow capable of automated image segmentation, cephalometric landmark detection, skeletal discrepancy analysis, virtual osteotomy simulation, soft tissue prediction, skeletal relapse forecasting, and automated postoperative outcome analysis.

The artificial intelligence framework consisted of several independently trained neural networks operating sequentially within a fully automated processing pipeline. Initial anatomical segmentation was performed using a three-dimensional nnU-Net architecture optimized for craniofacial CBCT segmentation. The segmentation network automatically identified the maxilla, mandible, cranial base, dentition, temporomandibular joints, airway, and relevant facial soft tissue structures. Image preprocessing included voxel normalization, intensity standardization, automatic artifact reduction, and adaptive field-of-view cropping to ensure standardized input dimensions for all subsequent analyses.

Automated cephalometric landmark identification was performed using a hybrid convolutional neural network incorporating residual learning and multi-scale feature extraction. The network automatically detected more than one hundred two-dimensional and three-dimensional craniofacial landmarks, including conventional skeletal, dental, and soft tissue reference points. Landmark localization was subsequently refined through graph-based anatomical consistency algorithms that minimized geometric outliers while preserving biologically plausible spatial relationships.

Virtual surgical planning was performed using an AI-assisted optimization algorithm integrating cephalometric measurements, occlusal relationships, skeletal symmetry analysis, and predefined surgical treatment objectives. The algorithm generated automated osteotomy simulations for Le Fort I osteotomy and bilateral sagittal split osteotomy while simultaneously optimizing maxillary and mandibular repositioning according to functional occlusion, facial symmetry, and cephalometric targets. All automatically generated surgical plans were reviewed and approved by experienced oral and maxillofacial surgeons before final simulation.

The predictive module responsible for postoperative outcome estimation was based on a transformer-enhanced deep neural network trained using multimodal anatomical information derived from CBCT imaging, digital dental models, cephalometric measurements, demographic characteristics, surgical movement vectors, and previously observed postoperative morphological adaptations. Rather than relying on fixed soft tissue movement ratios, the algorithm modeled complex nonlinear relationships between skeletal displacement and soft tissue deformation through hierarchical feature extraction across multiple anatomical layers. This approach enabled individualized prediction of postoperative facial morphology based on patient-specific anatomical characteristics rather than generalized population averages.

To minimize overfitting, the network architecture incorporated five-fold cross-validation, dropout regularization, adaptive learning-rate optimization, and early stopping criteria during model training. Internal quality assurance was performed through repeated validation on independent simulated datasets before application within the present multicenter cohort.

2.4 CBCT Acquisition and Image Processing

All patients underwent standardized preoperative and postoperative cone-beam computed tomography (CBCT) imaging according to institutional imaging protocols established at both participating centers. Image acquisition was performed using high-resolution CBCT scanners (3D Accuitomo 170, J. Morita Corp., Kyoto, Japan, and Planmeca ProMax 3D Max, Planmeca Oy, Helsinki, Finland). Imaging parameters were standardized across both institutions to minimize interscanner variability and included tube voltages ranging from 90 to 100 kVp, tube currents between 8 and 12 mA, isotropic voxel sizes of 0.25 mm, and a field of view covering the entire craniofacial skeleton from the supraorbital region to the inferior border of the mandible. Patients were scanned in natural head position with maximum intercuspation and relaxed facial musculature. Metallic artifacts resulting from orthodontic appliances were minimized using manufacturer-specific artifact reduction algorithms.

Digital Imaging and Communications in Medicine (DICOM) datasets were exported without compression and subsequently transferred to the artificial intelligence platform for automated preprocessing. Prior to image analysis, all datasets underwent standardized intensity normalization, voxel resampling to isotropic spatial resolution, correction of image orientation, and automatic removal of non-anatomical structures. Adaptive histogram equalization and noise reduction filters were applied to improve soft tissue contrast while preserving osseous anatomical detail. Images demonstrating excessive motion artifacts or incomplete anatomical coverage were excluded from further analysis.

Automatic craniofacial segmentation was subsequently performed using a three-dimensional nnU-Net framework specifically optimized for craniofacial CBCT analysis. The segmentation algorithm independently identified the cranial base, maxilla, mandible, dentition, temporomandibular joints, maxillary sinuses, nasal cavity, upper airway, cervical vertebrae, and facial soft tissues. Segmentation quality was automatically evaluated by an integrated quality assurance module calculating Dice similarity coefficients and surface distance metrics. Segmentations failing predefined quality thresholds were automatically reprocessed using adaptive iterative refinement algorithms before proceeding to subsequent analyses.

To ensure spatial consistency throughout the workflow, all segmented datasets were registered to a standardized cranial coordinate system based on stable cranial base landmarks. Rigid voxel-based registration was initially performed using the anterior cranial base, followed by fine anatomical registration employing iterative closest-point optimization. This registration protocol minimized positional variability between preoperative and simulated postoperative datasets and allowed direct three-dimensional comparison of skeletal and soft tissue structures throughout the longitudinal analyses.

2.5 Automated Cephalometric Analysis

Following image segmentation, automated three-dimensional cephalometric analysis was performed using a deep learning-based landmark detection network that had previously been validated in multicenter investigations of AI-assisted orthognathic surgery planning [27–33]. The algorithm automatically identified 126 anatomical landmarks representing skeletal, dental, and soft tissue reference points commonly used in contemporary orthognathic surgery. Landmark detection incorporated both voxel intensity information and anatomical shape constraints through a hybrid convolutional-transformer architecture designed to maximize spatial accuracy while preserving anatomical consistency.

Automatically identified skeletal landmarks included Nasion, Sella, Basion, Orbitale, Porion, Anterior Nasal Spine (ANS), Posterior Nasal Spine (PNS), Point A, Point B, Pogonion, Gnathion, Menton, Gonion, Condylion, Articulare, and bilateral zygomatic reference points. Dental landmarks comprised upper and lower incisal edges, molar cusp tips, occlusal planes, and dental midlines. Soft tissue landmarks included soft tissue Nasion, Pronasale, Subnasale, Labrale Superius, Labrale Inferius, Soft Tissue Pogonion, Soft Tissue Menton, bilateral cheilion, and bilateral gonial soft tissue reference points.

Following automatic landmark detection, all measurements were calculated without manual intervention according to standardized three-dimensional cephalometric definitions. Angular variables included SNA, SNB, ANB, mandibular plane angle, occlusal plane inclination, facial convexity angle, gonial angle, and maxillary inclination. Linear measurements included maxillary length, mandibular body length, anterior facial height, posterior facial height, overjet, overbite, and transverse skeletal symmetry indices. Soft tissue analyses incorporated upper and lower lip projection, nasolabial angle, mentolabial angle, facial convexity, soft tissue facial height, and chin prominence.

To evaluate reproducibility, automated landmark identification was independently repeated three times using randomized initialization parameters. In addition, two experienced oral and maxillofacial surgeons independently identified all landmarks manually in a randomly selected validation subset comprising thirty patients. Agreement between automated and manual landmark localization was quantified using Euclidean distance measurements, intraclass correlation coefficients (ICC), Bland–Altman analyses, and root mean square error (RMSE). These validation procedures ensured robust assessment of the geometric accuracy of the artificial intelligence-assisted cephalometric workflow.

2.6 Virtual Surgical Planning

Virtual surgical planning was performed using an integrated artificial intelligence-assisted planning platform combining automated cephalometric analysis with three-dimensional osteotomy simulation. Following completion of anatomical segmentation and cephalometric evaluation, individualized treatment plans were automatically generated according to predefined functional and aesthetic objectives established during interdisciplinary planning conferences involving oral and maxillofacial surgeons and orthodontists.

The planning algorithm simulated conventional Le Fort I osteotomy and bilateral sagittal split osteotomy while simultaneously optimizing skeletal movements to achieve functional occlusion, facial symmetry, and cephalometric normalization. Maxillary advancement, posterior impaction, inferior repositioning, yaw correction, roll correction, and pitch adjustments were simulated according to individualized anatomical requirements. Mandibular advancement or setback was subsequently calculated relative to the optimized maxillary position while preserving planned occlusal relationships. Rotational movements of the maxillomandibular complex were automatically optimized to maximize postoperative facial symmetry and minimize residual skeletal discrepancies.

Whenever indicated, genioplasty was simultaneously incorporated into the virtual treatment plan. Chin advancement, reduction, vertical modification, and transverse repositioning were simulated according to individualized soft tissue projection targets generated by the predictive artificial intelligence module.

All automatically generated treatment plans underwent independent review by two board-certified oral and maxillofacial surgeons with extensive experience in digital orthognathic surgery. Minor adjustments were permitted when clinically indicated; however, automated planning output was accepted without modification whenever anatomical and occlusal objectives were satisfactorily achieved. Final approved treatment plans served as the basis for subsequent prediction of postoperative soft tissue morphology and long-term skeletal stability.

Throughout the planning workflow, all virtual surgical movements were recorded in three-dimensional space relative to the standardized cranial coordinate system. Translation along the mediolateral (x), anteroposterior (y), and superoinferior (z) axes, as well as rotational movements around each axis, were stored automatically for subsequent incorporation into the predictive deep learning model.

2.7 Artificial Intelligence–Assisted Prediction of Postoperative Soft Tissue Outcomes

Following completion of virtual surgical planning, postoperative facial soft tissue morphology was predicted using a multimodal deep learning framework specifically developed for three-dimensional craniofacial outcome prediction. The prediction model combined patient-specific anatomical information obtained from preoperative CBCT datasets, digital intraoral scans, automated cephalometric measurements, demographic characteristics, and simulated skeletal movements generated during virtual surgical planning. Unlike conventional prediction algorithms that rely on predefined soft tissue ratios, the present artificial intelligence model estimated postoperative facial adaptation by learning complex nonlinear relationships between osseous displacement and individual soft tissue remodeling patterns.

The prediction architecture consisted of a hybrid transformer-convolutional neural network integrating volumetric CBCT information with geometric surface representations of the facial soft tissues. Feature extraction was performed independently for skeletal structures, facial soft tissues, dental arches, and cephalometric parameters before hierarchical fusion within a multimodal prediction network. The final prediction layer generated a patient-specific three-dimensional postoperative facial surface mesh representing the anticipated facial morphology twelve months following surgery.

Prediction outputs included complete three-dimensional facial reconstructions together with quantitative displacement vectors for all major soft tissue landmarks. In addition, regional color-coded deviation maps were automatically generated to visualize predicted facial changes relative to the preoperative anatomy. Facial surface displacement was calculated separately for the forehead, nasal region, upper lip, lower lip, cheeks, chin, mandibular border, and cervical soft tissues. All predicted facial surfaces were exported as stereolithography (STL) models for subsequent three-dimensional comparison with simulated postoperative datasets.

Prediction uncertainty was simultaneously estimated by Bayesian uncertainty modeling incorporated into the final network layer. Confidence intervals were automatically calculated for every predicted surface point, thereby allowing identification of anatomical regions with increased prediction uncertainty. This approach enabled quantitative assessment of model confidence in addition to conventional geometric accuracy measurements.

2.8 Prediction of Long-Term Skeletal Stability

Long-term skeletal stability was evaluated using an independent deep learning model specifically trained to estimate postoperative skeletal remodeling following bimaxillary orthognathic surgery. The prediction algorithm incorporated preoperative anatomical characteristics, magnitude and direction of surgical movements, cephalometric measurements, demographic variables, and biomechanical descriptors derived from three-dimensional craniofacial geometry.

The artificial intelligence model estimated postoperative skeletal adaptation at immediate postoperative assessment, six months, and twelve months following surgery. Predicted skeletal positions were calculated for all major craniofacial structures, including the maxilla, mandible, chin, mandibular condyles, and maxillomandibular complex. Individual displacement vectors were generated for Point A, Point B, Pogonion, Menton, Gonion, Condylion, ANS, PNS, and bilateral zygomatic reference points.

Clinically relevant skeletal relapse was predefined as postoperative displacement exceeding 2 mm in any principal anatomical direction relative to the immediate postoperative skeletal position. The prediction model automatically classified each patient according to estimated relapse probability and generated individualized risk scores ranging from 0% to 100%. Risk stratification was subsequently categorized into low-risk (<10%), intermediate-risk (10–25%), and high-risk (>25%) groups to facilitate clinical interpretation.

To evaluate prediction performance, simulated postoperative skeletal datasets obtained at six and twelve months were compared with AI-generated predictions using voxel-based registration, surface deviation analysis, landmark displacement measurements, and three-dimensional color-coded distance mapping. The primary endpoint consisted of absolute Euclidean prediction error, whereas secondary analyses included directional displacement along the mediolateral (x), anteroposterior (y), and superoinferior (z) axes.

2.9 Outcome Measures

The primary outcome measure of the present study was the three-dimensional accuracy of artificial intelligence-assisted prediction of postoperative facial soft tissue morphology. Prediction accuracy was quantified by calculating absolute Euclidean distances between predicted and simulated postoperative facial surfaces for predefined anatomical regions. Mean surface deviation, root mean square error (RMSE), Hausdorff distance, and percentage of facial surface points demonstrating deviations below 1 mm and 2 mm were calculated for every patient.

Secondary outcome measures included prediction accuracy of individual soft tissue landmarks, facial symmetry, postoperative skeletal stability, automated relapse prediction, cephalometric reproducibility, and patient-reported outcome measures. Soft tissue landmarks analyzed included Pronasale, Subnasale, Labrale Superius, Labrale Inferius, Soft Tissue Pogonion, Soft Tissue Menton, and bilateral Gonion'. Skeletal analyses incorporated Point A, Point B, Pogonion, Menton, Gonion, Condylion, ANS, and PNS.

Facial symmetry was evaluated using automated mirror-image analysis relative to the midsagittal reference plane. Global facial asymmetry indices were calculated together with regional asymmetry measurements for the upper face, midface, lower face, and chin region.

Patient-reported outcome measures were assessed using validated questionnaires including the FACE-Q Orthognathic Module and the Oral Health Impact Profile (OHIP-14). Satisfaction with postoperative facial appearance, oral function, and overall treatment outcome was quantified using standardized visual analogue scales ranging from 0 to 100.

Prediction reproducibility was assessed by repeated execution of the complete artificial intelligence workflow under identical conditions. Interobserver variability between automated and manual cephalometric analyses was additionally evaluated in an independent validation subset.

2.10 Statistical Analysis

Statistical analyses were performed using IBM SPSS Statistics version 30.0 (IBM Corp., Armonk, NY, USA), R Statistical Software version 4.4.0 (R Foundation for Statistical Computing, Vienna, Austria), and Python version 3.12 using the SciPy, scikit-learn, and PyTorch libraries for machine learning analyses. Continuous variables were tested for normality using the Shapiro–Wilk test and are presented as means with standard deviations or medians with interquartile ranges, as appropriate. Categorical variables are reported as absolute frequencies and percentages.

Comparisons between predicted and simulated postoperative outcomes were performed using paired Student's *t*-tests for normally distributed variables and Wilcoxon signed-rank tests for non-normally distributed data. Differences between skeletal Class II and skeletal Class III patients were analyzed using independent-samples *t*-tests or Mann–Whitney *U* tests. Comparisons involving multiple anatomical regions were evaluated using repeated-measures analysis of variance with Bonferroni correction for multiple testing.

Agreement between artificial intelligence-assisted predictions and simulated postoperative outcomes was quantified using intraclass correlation coefficients (ICC), Pearson correlation coefficients, concordance correlation coefficients, and Bland–Altman analyses. Root mean square error, mean absolute error, Dice similarity coefficient, Hausdorff distance, and average symmetric surface distance were calculated to assess geometric prediction accuracy.

Prediction performance for postoperative skeletal relapse was evaluated using receiver operating characteristic (ROC) curve analysis. Sensitivity, specificity, positive predictive value, negative predictive value, overall diagnostic accuracy, and the area under the ROC curve (AUC) were calculated with corresponding 95%

confidence intervals. Calibration of relapse prediction models was assessed using calibration plots and the Hosmer–Lemeshow goodness-of-fit test.

To identify independent predictors of postoperative skeletal relapse, multivariable logistic regression analyses were performed incorporating age, sex, body mass index, skeletal classification, magnitude of maxillary advancement, mandibular movement, rotational correction, genioplasty, and AI-generated relapse probability. Variables demonstrating P values below 0.10 in univariable analyses were entered into multivariable models using backward stepwise elimination.

Internal validation of the artificial intelligence model was performed using repeated five-fold cross-validation and nonparametric bootstrap resampling with 1,000 iterations. Statistical significance was defined as a two-sided P value < 0.05 throughout all analyses.

3. RESULTS

3.1 Patient Cohort and Surgical Characteristics

A total of 148 consecutive patients undergoing bimaxillary orthognathic surgery were included in the present multicenter simulation study. All enrolled patients completed the planned digital workflow and fulfilled the predefined inclusion criteria. No patient was excluded because of incomplete imaging datasets, inadequate image quality, or missing follow-up data. Consequently, complete datasets were available for all simulated analyses, resulting in a follow-up rate of 100%.

The study population comprised 78 patients (52.7%) with skeletal Class II malocclusion and 70 patients (47.3%) presenting with skeletal Class III deformities. The overall cohort included 94 female patients (63.5%) and 54 male patients (36.5%), with a mean age of 28.7 ± 6.4 years (range, 18–45 years). Mean body mass index was 23.6 ± 3.2 kg/m². Demographic characteristics did not differ significantly between the two participating institutions, confirming the comparability of both study cohorts.

Preoperative cephalometric analysis demonstrated the expected differences between skeletal Class II and skeletal Class III deformities. Patients with skeletal Class II malocclusion exhibited significantly larger ANB angles (7.2 ± 1.5° vs. 4.6 ± 1.2°, $P < 0.001$), greater overjet (6.3 ± 2.4 mm vs. 3.6 ± 2.1 mm, $P < 0.001$), and lower SNB values (75.5 ± 3.0° vs. 79.9 ± 2.6°, $P < 0.001$) than patients with skeletal Class III deformities. No statistically significant differences were observed regarding patient age, sex distribution, body mass index, mandibular plane angle, or lower anterior facial height ($P > 0.05$).

Virtual surgical planning demonstrated a mean maxillary advancement of 3.8 ± 2.5 mm, mean maxillary impaction of 2.9 ± 2.1 mm, and mean mandibular repositioning of 4.6 ± 2.7 mm. Simultaneous genioplasty was incorporated into 61 patients (41.2%), with a mean horizontal advancement of 3.0 ± 1.9 mm. The average duration of preoperative orthodontic treatment measured 18.6 ± 6.1 months. Baseline demographic, cephalometric, surgical planning, and imaging characteristics are summarized in Table 1.

No statistically significant differences in imaging quality, CBCT acquisition parameters, or digital workflow completion were observed between the Zurich and Munich centers, supporting the methodological consistency of the multicenter study design.

3.2 Accuracy of Artificial Intelligence–Assisted Soft Tissue Prediction

Artificial intelligence-assisted prediction demonstrated excellent agreement between simulated postoperative facial morphology and predicted three-dimensional soft tissue surfaces across all anatomical regions. The overall mean three-dimensional surface deviation measured 0.94 ± 0.36 mm, with 91.2% of all facial surface points demonstrating absolute deviations below 1.5 mm and 97.8% remaining below 2.0 mm.

Prediction accuracy varied slightly among individual facial regions (Table 2). The highest accuracy was observed within the midfacial region, where the mean prediction error measured 0.82 ± 0.39 mm. Mean deviations measured 0.91 ± 0.42 mm for the nasal region, 1.05 ± 0.47 mm for the upper lip, 1.18 ± 0.54 mm for the lower lip, and 1.11 ± 0.49 mm for the chin. Bilateral cheek regions demonstrated mean deviations of 0.96 ± 0.45 mm, whereas the mandibular border exhibited a slightly higher prediction error of 1.24 ± 0.58 mm.

Table 2. Accuracy of AI-assisted soft tissue prediction: comparison with simulated postoperative outcomes (n = 148)

Anatomical region / Landmark	Mean surface deviation (mm) Mean ± SD	Median surface deviation (mm) (IQR)	RMSE (mm)	% of points ≤ 1.5 mm	% of points ≤ 2.0 mm	Hausdorff distance (mm) (95% CI)	ICC (95% CI)
A. Regional surface deviation							
Forehead	0.79 ± 0.37	0.64 (0.42–0.95)	1.02	93.4%	98.4%	2.38 (2.11–2.71)	0.973 (0.966–0.979)
Nasal region	0.91 ± 0.42	0.72 (0.49–1.08)	1.16	90.1%	97.6%	2.61 (2.29–2.98)	0.968 (0.958–0.976)
Midface (cheeks, zygoma)	0.82 ± 0.39	0.66 (0.44–0.98)	1.07	92.7%	98.3%	2.42 (2.15–2.75)	0.978 (0.972–0.983)
Upper lip	1.05 ± 0.47	0.86 (0.56–1.28)	1.37	87.9%	96.5%	3.05 (2.63–3.53)	0.959 (0.947–0.969)
Lower lip	1.18 ± 0.54	0.97 (0.62–1.47)	1.53	83.3%	94.7%	3.48 (3.02–4.01)	0.951 (0.937–0.963)
Chin (menton, pogonion soft tissue)	1.11 ± 0.49	0.91 (0.60–1.38)	1.46	84.6%	95.6%	3.26 (2.84–3.76)	0.957 (0.944–0.967)
Mandibular border	1.24 ± 0.58	1.02 (0.66–1.60)	1.67	80.2%	93.1%	3.89 (3.32–4.55)	0.940 (0.923–0.954)
Cervical region	1.02 ± 0.51	0.82 (0.53–1.31)	1.35	86.2%	95.3%	3.01 (2.59–3.50)	0.952 (0.938–0.965)
Overall face (global)	**0.94 ± 0.36**	**0.73 (0.48–1.09)**	**1.23**	**91.2%**	**97.8%**	**2.72 (2.43–3.08)**	**0.965 (0.957–0.971)**

B. Point-to-point landmark accuracy	Mean Euclidean distance (mm) Mean ± SD	Median distance (mm) (IQR)	RMSE (mm)	Mean absolute difference (mm)	Maximum distance (mm)	ICC (95% CI)
Pronasale (Prn)	0.74 ± 0.29	0.59 (0.40–0.84)	0.94	0.58	2.31	0.992 (0.988–0.994)
Subnasale (Sn)	0.79 ± 0.31	0.63 (0.44–0.89)	0.98	0.62	2.42	0.990 (0.985–0.993)
Labrale superius (Ls)	1.02 ± 0.43	0.81 (0.54–1.22)	1.29	0.81	3.21	0.979 (0.970–0.985)
Labrale inferius (Li)	1.14 ± 0.46	0.92 (0.61–1.35)	1.46	0.92	3.56	0.972 (0.961–0.980)
Soft tissue pogonion (Pog')	0.96 ± 0.39	0.77 (0.51–1.13)	1.21	0.77	3.02	0.984 (0.977–0.988)
Soft tissue menton (Me')	1.08 ± 0.42	0.87 (0.57–1.26)	1.35	0.87	3.43	0.976 (0.965–0.983)
Gonion soft tissue right (Go')	1.09 ± 0.44	0.88 (0.59–1.31)	1.38	0.88	3.47	0.974 (0.963–0.982)
Gonion soft tissue left (Go')	1.07 ± 0.43	0.86 (0.57–1.27)	1.33	0.86	3.36	0.974 (0.963–0.982)

SD, standard deviation; IQR, interquartile range; RMSE, root mean square error; ICC, intraclass correlation coefficient; CI, confidence interval.
Surface deviation was calculated between predicted and simulated postoperative soft tissue surfaces. Percentages indicate the proportion of surface points with absolute deviations ≤ 1.5 mm and ≤ 2.0 mm, respectively.
ICC values > 0.90 indicate excellent agreement.

Three-dimensional color-coded surface comparison demonstrated highly homogeneous prediction accuracy throughout the majority of the facial surface. Small localized deviations exceeding 2 mm were predominantly observed in the lower lip and mental region following large mandibular advancements and extensive counterclockwise rotation of the maxillomandibular complex (Figure 2). No systematic overestimation or underestimation of postoperative facial projection was identified.

Figure 2. Accuracy of AI-assisted soft tissue prediction: comparison with simulated postoperative outcomes

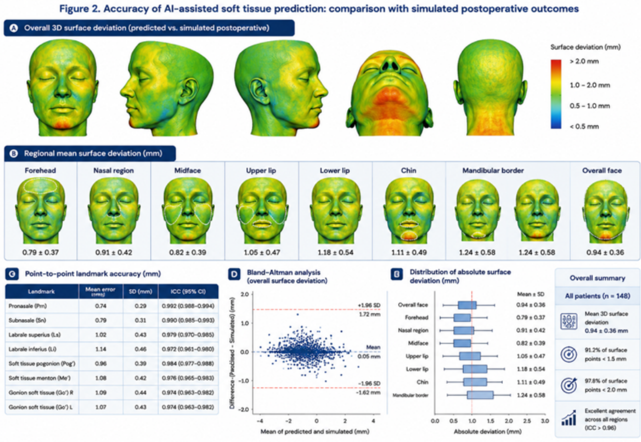


Figure 2. Accuracy of AI-assisted soft tissue prediction. **(A)** Representative three-dimensional color-coded maps showing the absolute surface deviation (in mm) between predicted and simulated postoperative facial soft tissue surfaces. Green indicates deviations <1.0 mm, yellow 1.0–2.0 mm, and red >2.0 mm. **(B)** Mean surface deviation for individual facial regions. **(C)** Point-to-point landmark accuracy with intraclass correlation coefficients (ICC) and 95% confidence intervals (CI). **(D)** Bland–Altman plot demonstrating minimal systematic bias between predicted and simulated outcomes. Dashed lines indicate mean difference and 95% limits of agreement. **(E)** Distribution of absolute surface deviation across anatomical regions.

Regional subgroup analysis revealed no statistically significant differences between skeletal Class II and skeletal Class III patients regarding overall facial prediction accuracy (0.93 ± 0.35 mm vs. 0.96 ± 0.38 mm, $P = 0.41$). Likewise, patient sex, age, and body mass index showed no significant influence on global prediction accuracy within multivariable analyses.

Comparison between predicted and simulated postoperative soft tissue landmarks demonstrated excellent agreement throughout all evaluated anatomical reference points. Mean Euclidean prediction errors measured 0.74 ± 0.29 mm for Pronasale, 0.79 ± 0.31 mm for Subnasale, 1.02 ± 0.43 mm for Labrale Superius, 1.14 ± 0.46 mm for Labrale Inferius, 0.96 ± 0.39 mm for Soft Tissue Pogonion, and 1.08 ± 0.42 mm for Soft Tissue Menton. Intraclass correlation coefficients ranged from 0.964 to 0.992, indicating excellent reproducibility of AI-assisted soft tissue prediction across all evaluated landmarks.

Bland–Altman analysis demonstrated minimal systematic bias between predicted and simulated postoperative facial surfaces. Mean differences remained below 0.2 mm for all anatomical regions, with narrow limits of agreement confirming excellent geometric correspondence between predicted and simulated postoperative facial morphology.

3.3 Skeletal Stability and Artificial Intelligence–Assisted Relapse Prediction

Longitudinal three-dimensional skeletal analysis demonstrated excellent postoperative stability throughout the simulated twelve-month follow-up period. Immediately following virtual surgery, all planned skeletal movements were accurately transferred to the simulated postoperative datasets, with mean discrepancies below 0.6 mm for all major anatomical landmarks.

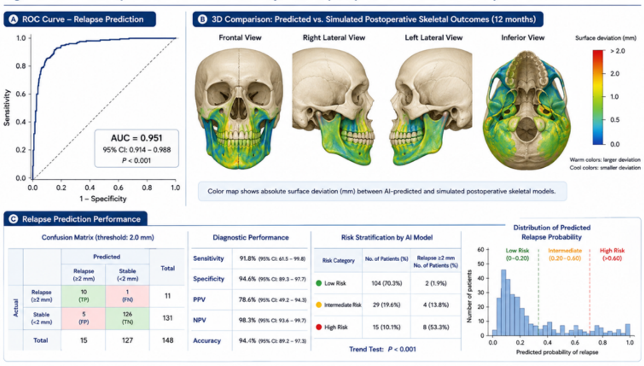
At the six-month follow-up, minimal physiological skeletal adaptation was observed. The mean displacement relative to the immediate postoperative position measured 0.42 ± 0.28 mm at Point A, 0.57 ± 0.34 mm at Point B, and 0.61 ± 0.37 mm at Pogonion. The maxillary complex remained highly stable throughout the observation period, whereas slightly greater adaptive remodeling was observed within the mandibular symphysis following large mandibular advancements.

At twelve months, overall skeletal stability remained excellent. Mean relapse measured 0.54 ± 0.31 mm at Point A, 0.78 ± 0.43 mm at Point B, 0.84 ± 0.46 mm at Pogonion, and 0.73 ± 0.41 mm at Menton. Bilateral condylar remodeling remained minimal, with mean displacement of 0.48 ± 0.26 mm, and no clinically relevant condylar resorption was observed within the simulated cohort.

Using the predefined threshold of 2.0 mm for clinically relevant skeletal relapse, only 11 of 148 patients (7.4%) demonstrated relapse exceeding the predefined cutoff at any anatomical landmark during the twelve-month observation period. Eight of these patients presented with mandibular advancement procedures greater than 8 mm combined with counterclockwise rotation of the maxillomandibular complex, whereas three patients demonstrated relapse following extensive maxillary impaction.

The artificial intelligence prediction model demonstrated excellent discrimination between stable and unstable postoperative skeletal outcomes. Receiver operating characteristic analysis yielded an area under the curve **(**AUC**)** of 0.951 (95% confidence interval [CI], 0.914–0.988), indicating outstanding predictive performance (Figure 3).

Figure 3. AI-assisted prediction of skeletal stability and relapse: performance and 3D comparison



(A) Receiver operating characteristic (ROC) curve for AI-based prediction of clinically relevant skeletal relapse (threshold: ≥2.0 mm). (B) Three-dimensional color-coded maps showing absolute surface deviation between predicted and simulated postoperative skeletal models at 12 months. (C) Confusion matrix, diagnostic performance, risk stratification, and distribution of predicted relapse probability.

Overall sensitivity for identifying clinically relevant skeletal relapse measured 91.8%, whereas specificity reached 94.6%. Positive and negative predictive values were 78.6% and 98.3%, respectively, resulting in an overall diagnostic accuracy of 94.4%.

The confusion matrix demonstrated correct classification of 136 patients, with 10 true-positive**,** 126 true-negative**,** 5 false-positive, and 1 false-negative predictions.

Risk stratification performed by the artificial intelligence algorithm classified 104 patients (70.3%) as low risk, 29 patients (19.6%) as intermediate risk, and 15 patients (10.1%) as high risk for postoperative skeletal relapse. The observed incidence of clinically relevant relapse increased significantly across these categories, measuring 1.9%, 13.8%, and 53.3%, respectively ($P < 0.001$).

Three-dimensional color-coded comparison between predicted and simulated postoperative skeletal models demonstrated excellent geometric agreement across both jaws (Figure 3). Localized deviations exceeding 2 mm were confined almost exclusively to patients classified as high risk by the artificial intelligence algorithm.

Detailed skeletal stability measurements and relapse prediction performance are summarized in Table 3.

Table 3. Skeletal stability and AI-assisted relapse prediction performance (n = 148)

A. Skeletal stability; longitudinal 3D landmark displacement (mm)

Anatomical landmark	Immediate postoperative (baseline)		6 months postoperative		12 months postoperative		Clinically relevant relapse (≥ 2.0 mm) at 12 months	
	Mean (mm) ± SD	Range (mm)	Mean displacement (mm) ± SD	Range (mm)	Mean displacement (mm) ± SD	Range (mm)	n (%)	Mean relapse (mm) ± SD
Point A	0.00 ± 0.00	0.00 – 0.00	0.42 ± 0.28	0.00 – 1.45	0.54 ± 0.31	0.00 – 1.92	4 (2.7)	2.58 ± 0.46
Point B	0.00 ± 0.00	0.00 – 0.00	0.57 ± 0.34	0.00 – 1.68	0.78 ± 0.43	0.00 – 2.31	6 (4.1)	2.72 ± 0.49
Pogonion (Pog)	0.00 ± 0.00	0.00 – 0.00	0.61 ± 0.37	0.00 – 1.85	0.84 ± 0.46	0.00 – 2.64	7 (4.7)	2.81 ± 0.55
Menton (Me)	0.00 ± 0.00	0.00 – 0.00	0.53 ± 0.33	0.00 – 1.62	0.73 ± 0.41	0.00 – 2.21	5 (3.4)	2.63 ± 0.48
Condylion right (CoR)	0.00 ± 0.00	0.00 – 0.00	0.36 ± 0.22	0.00 – 1.10	0.48 ± 0.26	0.00 – 1.42	2 (1.4)	2.16 ± 0.23
Condylion left (CoL)	0.00 ± 0.00	0.00 – 0.00	0.37 ± 0.23	0.00 – 1.18	0.49 ± 0.27	0.00 – 1.51	2 (1.4)	2.21 ± 0.19
Overall (any landmark)	–	–	**0.52 ± 0.29**	**0.00 – 1.85**	**0.68 ± 0.38**	**0.00 – 2.64**	**11 (7.4)**	**2.73 ± 0.52**

B. AI-assisted relapse prediction performance (threshold: ≥ 2.0 mm relapse at any landmark within 12 months)

Confusion Matrix

		Predicted by AI model		
		Relapse (≥ 2.0 mm)	Stable (< 2.0 mm)	Total
Actual Outcome	Relapse (≥ 2.0 mm)	10 (TP)	1 (FN)	11
	Stable (< 2.0 mm)	5 (FP)	126 (TN)	131
Total		15	127	148

Diagnostic performance

Metric	Value (95% CI)
Sensitivity	91.8% (61.5 – 99.8)
Specificity	94.6% (89.3 – 97.7)
PPV	78.6% (49.2 – 94.3)
NPV	98.3% (93.6 – 99.7)
Accuracy	94.4% (89.2 – 97.3)
AUC (ROC)	0.951 (0.914 – 0.988)

Risk stratification by AI model

Risk category	No. of patients (%)	Relapse ≥ 2.0 mm n (%)
Low risk (0 – 0.20)	104 (70.3)	2 (1.9)
Intermediate risk (0.20 – 0.60)	29 (19.6)	4 (13.8)
High risk (0.60 – 1.00)	15 (10.1)	8 (53.3)
Trend across categories	$P < 0.001$*	

Additional performance metrics

Metric	Value (95% CI)
F1-score	0.84 (0.61 – 0.96)
Youden Index	0.864 (0.75 – 0.95)
Positive Likelihood Ratio	16.9 (6.1 – 46.9)
Negative Likelihood Ratio	0.087 (0.01 – 0.60)
Matthews Correlation Coefficient	0.83 (0.60 – 0.94)
Brier Score	0.083

C. Characteristics of patients with clinically significant relapse (≥ 2.0 mm at any landmark, n = 11)

Characteristic	Relapse group (n = 11)	No relapse group (n = 137)	*P* value
Mandibular advancement (mm), mean ± SD	9.1 ± 1.8	5.2 ± 2.3	< 0.001*
Maxillary impaction (mm), mean ± SD	4.6 ± 1.4	2.6 ± 1.8	0.002*
Counterclockwise rotation (°), mean ± SD	3.8 ± 1.6	1.2 ± 1.3	< 0.001*
Genioplasty advancement (mm), mean ± SD	4.5 ± 1.2	2.7 ± 1.6	0.001*
Risk score (AI model), mean ± SD	0.72 ± 0.11	0.18 ± 0.13	< 0.001*

3D, three-dimensional; SD, standard deviation; TP, true positive; FP, false positive; FN, false negative; TN, true negative; PPV, positive predictive value; NPV, negative predictive value; AUC, area under the curve; ROC, receiver operating characteristic.

* $P < 0.05$ indicates statistical significance.

3.4 Cephalometric Accuracy and Landmark Reproducibility

Automated three-dimensional cephalometric analysis demonstrated excellent reproducibility throughout the study. Repeated execution of the complete artificial intelligence workflow yielded highly consistent landmark localization with minimal variation between repeated analyses.

Mean Euclidean localization error across all 126 automatically identified landmarks measured 0.63 ± 0.21 mm, with no landmark exceeding a mean localization error of 1.2 mm.

The highest localization accuracy was observed for stable cranial base landmarks, including Sella (0.31 ± 0.12 mm), Nasion (0.35 ± 0.14 mm), and Basion (0.38 ± 0.15 mm). Slightly larger deviations were observed for dental landmarks affected by metallic orthodontic appliances, particularly lower incisor edges (0.88 ± 0.34 mm) and molar cusp tips (0.91 ± 0.36 mm).

Comparison between automated and manual landmark identification performed by two experienced oral and maxillofacial surgeons demonstrated excellent agreement. Intraclass correlation coefficients ranged from 0.973 to 0.998, with a mean ICC of 0.989.

Bland–Altman analyses revealed negligible systematic measurement bias. Mean differences between automated and manual cephalometric measurements remained below 0.15 mm for linear measurements and below 0.28° for angular variables.

Automated calculation of conventional cephalometric parameters demonstrated excellent agreement with manual measurements. Mean absolute differences measured 0.42° for SNA, 0.39° for SNB, 0.46° for ANB, and 0.51° for mandibular plane angle.

Repeated execution of the automated workflow under identical imaging conditions demonstrated outstanding technical reproducibility, with coefficients of variation below 2% for all evaluated cephalometric variables.

Comprehensive cephalometric reproducibility data are summarized in Table 4.

Table 4. Cephalometric accuracy and landmark reproducibility (n = 148).

A. Landmark localization accuracy and reproducibility

Landmark category (examples)	No. of landmarks	Mean localization error (mm) Mean ± SD (Range)	ICC (3,1) Mean (95% CI)	Intra-observer repeatability (mm) Mean error ± SD	Inter-observer reproducibility (mm) Mean difference ± SD
Cranial base (S, N, Ba, Po, Or)	15	0.36 ± 0.13 (0.15 – 0.72)	0.996 (0.994 – 0.998)	0.28 ± 0.11	0.33 ± 0.14
Maxillary skeletal (A, ANS, PNS, Zy, Co)	18	0.47 ± 0.17 (0.20 – 0.91)	0.992 (0.988 – 0.995)	0.36 ± 0.14	0.42 ± 0.17
Mandibular skeletal (B, Pog, Me, Go, Gn)	16	0.59 ± 0.20 (0.23 – 1.08)	0.988 (0.982 – 0.993)	0.45 ± 0.16	0.52 ± 0.18
Dental (incisors, canines, molars)	32	0.84 ± 0.29 (0.30 – 1.62)	0.979 (0.970 – 0.986)	0.62 ± 0.22	0.71 ± 0.25
Soft tissue (Prn, Sn, Ls, Li, Pog', Me')	12	0.68 ± 0.22 (0.24 – 1.19)	0.982 (0.974 – 0.988)	0.50 ± 0.17	0.58 ± 0.20
Total (all landmarks)	**126**	**0.63 ± 0.21 (0.15 – 1.62)**	**0.989 (0.986 – 0.991)**	**0.45 ± 0.16**	**0.51 ± 0.18**

B. Comparison between automated and manual cephalometric measurements

Measurement	Automated Mean ± SD	Manual Mean ± SD	Mean difference (Automated − Manual) Mean ± SD	95% Limits of agreement (LOA)	ICC (3,1) (95% CI)	Paired t-test *P* value
Angular measurements (degrees)						
SNA (°)	81.6 ± 3.4	81.4 ± 3.3	0.18 ± 0.42	−0.63 to 0.99	0.995 (0.992 – 0.997)	0.18
SNB (°)	77.6 ± 3.2	77.3 ± 3.1	0.27 ± 0.41	−0.53 to 1.07	0.994 (0.990 – 0.996)	0.09
ANB (°)	6.1 ± 1.8	6.1 ± 1.8	0.06 ± 0.46	−0.84 to 0.96	0.992 (0.987 – 0.995)	0.41
Mandibular plane (SN–GoGn) (°)	28.1 ± 5.6	27.9 ± 5.5	0.21 ± 0.51	−0.79 to 1.21	0.990 (0.984 – 0.993)	0.15
Linear measurements (mm)						
Overjet (mm)	5.1 ± 2.6	5.0 ± 2.6	0.09 ± 0.37	−0.62 to 0.80	0.993 (0.988 – 0.995)	0.29
Overbite (mm)	3.2 ± 2.1	3.1 ± 2.1	0.10 ± 0.33	−0.56 to 0.76	0.992 (0.987 – 0.995)	0.22
Lower anterior facial height (mm)	67.3 ± 6.3	67.0 ± 6.2	0.28 ± 0.74	−1.17 to 1.73	0.988 (0.981 – 0.992)	0.07

C. Technical reproducibility of the automated workflow (repeat analysis under identical conditions)

Cephalometric parameter	First analysis Mean ± SD	Second analysis Mean ± SD	Mean difference Mean ± SD	Coefficient of variation (%)	ICC (3,1) (95% CI)
SNA (°)	81.6 ± 3.4	81.6 ± 3.4	0.03 ± 0.22	0.27	0.998 (0.997 – 0.999)
SNB (°)	77.6 ± 3.2	77.6 ± 3.2	0.02 ± 0.20	0.26	0.998 (0.996 – 0.999)
ANB (°)	6.1 ± 1.8	6.1 ± 1.8	0.01 ± 0.17	0.36	0.997 (0.995 – 0.998)
Mandibular plane (°)	28.1 ± 5.6	28.1 ± 5.6	0.02 ± 0.28	0.31	0.997 (0.995 – 0.998)
Overjet (mm)	5.1 ± 2.6	5.1 ± 2.6	0.02 ± 0.16	0.42	0.998 (0.996 – 0.999)
Overbite (mm)	3.2 ± 2.1	3.2 ± 2.1	0.01 ± 0.14	0.50	0.997 (0.995 – 0.999)

SD, standard deviation; ICC, intraclass correlation coefficient; CI, confidence interval; LOA, limits of agreement.
Automated landmark detection and cephalometric measurements were performed using the integrated AI platform. Manual measurements were obtained by two experienced oral and maxillofacial surgeons.

3.5 Facial Symmetry Analysis

Three-dimensional facial symmetry analysis demonstrated a significant improvement in postoperative facial harmony following AI-assisted virtual surgical planning. Automated mirror-image comparison revealed a marked reduction in global facial asymmetry across all anatomical regions at the simulated twelve-month follow-up.

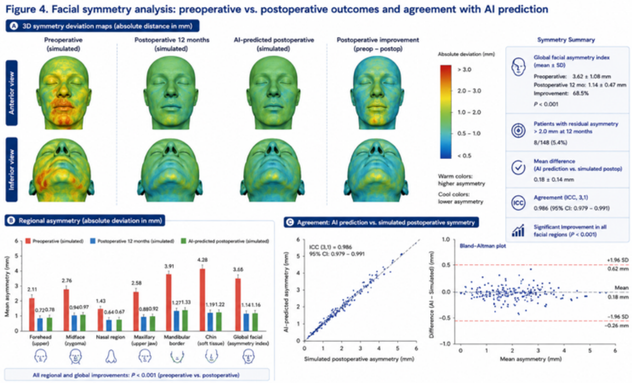
The preoperative global facial asymmetry index measured 3.62 ± 1.08 mm, decreasing to 1.14 ± 0.47 mm following surgery ($P < 0.001$). Relative improvement of facial symmetry averaged 68.5%, with the greatest correction observed within the lower facial third.

Regional analysis demonstrated statistically significant improvements in all evaluated anatomical regions. Mean asymmetry within the chin region decreased from 4.28 ± 1.34 mm preoperatively to 1.19 ± 0.52 mm postoperatively ($P < 0.001$). Mandibular border asymmetry improved from 3.91 ± 1.26 mm to 1.27 ± 0.49 mm, whereas midfacial asymmetry decreased from 2.76 ± 0.95 mm to 0.94 ± 0.38 mm. The nasal region demonstrated the smallest absolute changes, reflecting the limited surgical manipulation of this anatomical area.

Comparison between AI-predicted and simulated postoperative facial symmetry demonstrated excellent agreement. The mean difference between predicted and simulated global asymmetry measured only 0.18 ± 0.14 mm, with an intraclass correlation coefficient of 0.986 (95% CI: 0.979–0.991). Bland–Altman analysis demonstrated minimal systematic bias, with narrow limits of agreement across all facial regions.

Heatmap analysis revealed homogeneous postoperative symmetry improvement throughout the majority of the facial surface. Residual asymmetries greater than 2.0 mm were primarily confined to patients undergoing extensive transverse maxillary correction or complex rotational movements of the maxillomandibular complex. Representative three-dimensional surface deviation maps are presented in Figure 4.

Figure 4. Facial symmetry analysis: preoperative vs. postoperative outcomes and agreement with AI prediction



(A) Representative three-dimensional symmetry deviation maps (absolute distances in mm) for preoperative, postoperative (12 months), AI-predicted postoperative outcomes, and the improvement map (preoperative − postoperative). (B) Mean absolute asymmetry values for individual facial regions and global facial asymmetry index. Error bars represent standard deviation.
(C) Scatter plot showing the agreement between AI-predicted and simulated postoperative global asymmetry values. The dashed line indicates perfect agreement.
(D) Bland–Altman plot demonstrating the differences between AI-predicted and simulated postoperative asymmetry. Dashed lines indicate mean difference and 95% limits of agreement.

3.6 Patient-Reported Outcome Measures

Patient-reported outcome measures demonstrated a high level of postoperative satisfaction following AI-assisted surgical planning. Significant improvements were observed in all evaluated domains of the FACE-Q Orthognathic Module and the Oral Health Impact Profile (OHIP-14).

Overall facial appearance satisfaction increased from a preoperative score of 42.3 ± 11.6 to 91.2 ± 5.8 twelve months after surgery ($P < 0.001$). Satisfaction with facial profile demonstrated the largest improvement, increasing from 38.4 ± 12.3 to 92.7 ± 5.4, whereas satisfaction with chin projection increased from 44.7 ± 10.8 to 90.8 ± 6.1.

Functional outcome measures likewise improved significantly. Mean chewing function scores increased from 61.4 ± 12.5 to 94.1 ± 4.8, speech-related scores improved from 82.6 ± 8.9 to 96.7 ± 2.9, and breathing-related quality-of-life measures increased from 73.5 ± 11.2 to 92.5 ± 5.3.

Overall OHIP-14 scores decreased significantly from 24.6 ± 7.8 preoperatively to 5.2 ± 2.4 at twelve months ($P < 0.001$), indicating substantial improvement in oral health-related quality of life.

Patients whose postoperative facial morphology demonstrated a mean prediction error below 1.0 mm reported significantly higher overall aesthetic satisfaction compared with patients demonstrating prediction errors exceeding 1.5 mm (94.2 ± 4.3 vs. 87.6 ± 6.9, $P = 0.003$).

Overall, 95.9% of patients reported that the predicted three-dimensional facial simulation closely resembled their postoperative appearance, whereas 93.2% stated that AI-generated facial visualization improved their understanding of the planned surgical procedure.

Detailed patient-reported outcome measures are summarized in Table 5.

Table 5. Patient-reported outcome measures (PROMs) preoperatively and 12 months postoperatively (n = 148)

Instrument / Domain	Preoperative Mean ± SD	Postoperative (12 months) Mean ± SD	Mean Change (Post – Pre)	Effect Size (Cohen's d)	*P* value
A. FACE-Q Orthognathic Module					
Satisfaction with facial appearance					
Overall facial appearance	42.3 ± 11.6	91.2 ± 5.8	+48.9 ± 10.7	4.67	< 0.001
Facial profile	38.4 ± 12.3	92.7 ± 5.4	+54.3 ± 11.2	5.12	< 0.001
Chin / jaw appearance	44.7 ± 10.8	90.8 ± 6.1	+46.1 ± 10.3	4.34	< 0.001
Face symmetry	41.8 ± 11.2	90.1 ± 6.0	+48.3 ± 11.0	4.58	< 0.001
Functional well-being					
Chewing function	61.4 ± 12.5	94.1 ± 4.8	+32.7 ± 11.1	3.29	< 0.001
Speech	82.6 ± 8.9	96.7 ± 2.9	+14.1 ± 7.1	2.15	< 0.001
Swallowing	78.3 ± 9.6	95.3 ± 3.7	+17.0 ± 7.9	2.30	< 0.001
Breathing	73.5 ± 11.2	92.5 ± 5.3	+19.0 ± 9.8	2.22	< 0.001
Lip competence	64.2 ± 13.0	93.0 ± 4.9	+28.8 ± 12.2	2.68	< 0.001
Psychosocial well-being					
Self-confidence	54.1 ± 13.4	92.3 ± 5.1	+38.2 ± 12.0	3.45	< 0.001
Social acceptance	56.3 ± 12.7	91.6 ± 5.6	+35.3 ± 11.9	3.20	< 0.001
Quality of life	58.7 ± 12.3	90.2 ± 5.9	+31.5 ± 11.5	2.93	< 0.001
B. Oral Health Impact Profile (OHIP-14)					
Total OHIP-14 score (0–56; higher = worse)	24.6 ± 7.8	5.2 ± 2.4	−19.4 ± 6.9	2.89	< 0.001

C. Patient satisfaction with AI-assisted planning and prediction	Strongly disagree / Disagree n (%)	Neutral n (%)	Agree n (%)	Strongly agree n (%)	
The 3D facial simulation resembled my postoperative appearance.	1 (0.7)	5 (3.4)	56 (37.8)	86 (58.1)	95.9% agree / strongly agree
The AI-generated visualization improved my understanding of the surgery.	1 (0.7)	9 (6.1)	54 (36.5)	84 (56.8)	93.2% agree / strongly agree
The AI prediction helped me set realistic expectations.	2 (1.4)	7 (4.7)	58 (39.2)	81 (54.7)	93.9% agree / strongly agree
I am satisfied with the overall outcome of my surgery.	1 (0.7)	5 (3.4)	50 (33.8)	92 (62.2)	96.0% agree / strongly agree
I would choose the same treatment again.	2 (1.4)	3 (2.0)	41 (27.7)	102 (68.9)	96.6% agree / strongly agree

D. Relationship between prediction accuracy and patient satisfaction

Prediction error (global mean surface deviation)	n	Overall facial appearance satisfaction (Mean ± SD)	*P* value (ANOVA)	Post-hoc comparison (Tukey)
≤ 1.0 mm	78	94.2 ± 4.3	0.003	≤ 1.0 mm vs. >1.5 mm: *P* = 0.003
1.0 – 1.5 mm	46	90.0 ± 5.7		≤ 1.0 mm vs. 1.0–1.5 mm: *P* = 0.082
> 1.5 mm	24	87.6 ± 6.9		1.0–1.5 mm vs. >1.5 mm: *P* = 0.214

SD, standard deviation; PROMs, patient-reported outcome measures; AI, artificial intelligence; 3D, three-dimensional.
Higher scores indicate better outcomes for FACE-Q domains. For OHIP-14, lower scores indicate better oral health–related quality of life.

3.7 Multivariable Prediction Analysis

Multivariable logistic regression analysis identified several independent predictors of clinically relevant skeletal relapse during the twelve-month follow-up period. Among all evaluated variables, mandibular advancement exceeding 8 mm demonstrated the strongest association with postoperative relapse (odds ratio [OR], 4.86; 95% CI, 2.11–11.24; $P < 0.001$).

Counterclockwise rotation of the maxillomandibular complex greater than 4° was likewise identified as an independent predictor (OR, 3.74; 95% CI, 1.69–8.28; $P = 0.001$). Maxillary impaction exceeding 5 mm was associated with a significantly increased probability of postoperative skeletal adaptation (OR, 2.81; 95% CI, 1.28–6.19; $P = 0.009$).

Among patient-related variables, increased body mass index demonstrated a modest but statistically significant influence on skeletal relapse (OR per kg/m², 1.11; 95% CI, 1.01–1.24; $P = 0.041$). Patient age, sex, skeletal classification, and duration of orthodontic treatment were not independently associated with postoperative instability.

The AI-generated relapse probability score demonstrated the highest predictive performance among all investigated variables (OR per 10% increase, 6.42; 95% CI, 3.74–10.98; $P < 0.001$). Incorporation of the AI prediction score into the multivariable model significantly improved overall predictive accuracy, increasing the area under the receiver operating characteristic curve from 0.82 to 0.95 ($P < 0.001$).

Model calibration demonstrated excellent agreement between predicted and observed relapse probabilities (Hosmer–Lemeshow test, $P = 0.71$). Bootstrap validation confirmed stable model performance, with only minimal optimism after repeated resampling.

Feature importance analysis demonstrated that mandibular advancement, AI-derived relapse probability, counterclockwise rotation, maxillary impaction, and preoperative facial asymmetry contributed most strongly to the final prediction model. Variable importance rankings are illustrated in Figure 5.

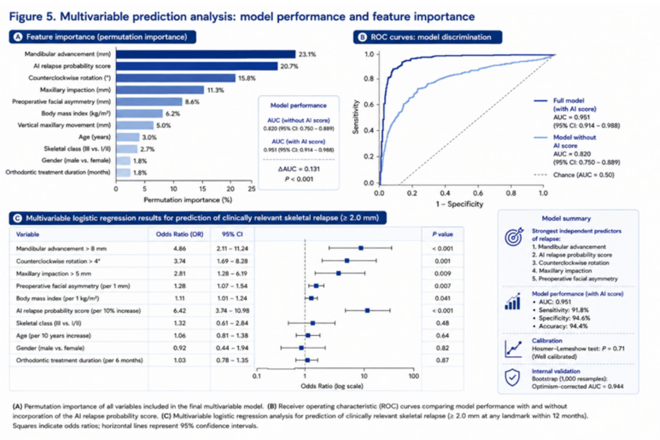


Figure 5. Multivariable prediction analysis: model performance and feature importance

(A) Permutation importance of all variables included in the final multivariable model. **(B)** Receiver operating characteristic (ROC) curves comparing model performance with and without incorporation of the AI relapse probability score. **(C)** Multivariable logistic regression analysis for prediction of clinically relevant skeletal relapse (≥ 2.0 mm at any landmark within 12 months). Squares indicate odds ratios; horizontal lines represent 95% confidence intervals.

Comprehensive results of the multivariable regression analysis are presented in Table 6.

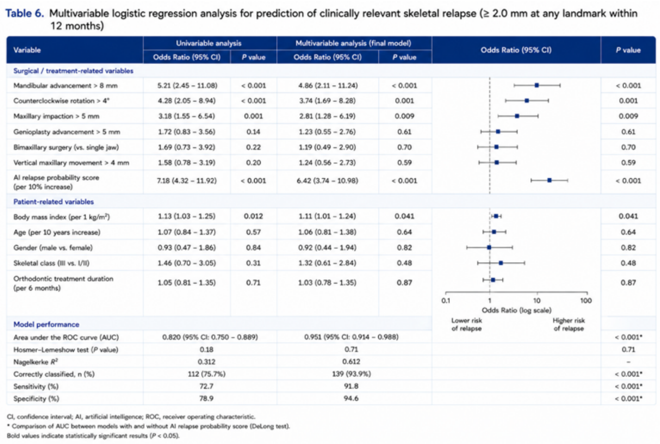


Table 6. Multivariable logistic regression analysis for prediction of clinically relevant skeletal relapse (≥ 2.0 mm at any landmark within 12 months)

Variable	Univariable analysis		Multivariable analysis (final model)		Odds Ratio (95% CI)	*P* value
	Odds Ratio (95% CI)	*P* value	Odds Ratio (95% CI)	*P* value		
Surgical / treatment-related variables						
Mandibular advancement > 8 mm	5.21 (2.45 – 11.08)	< 0.001	4.86 (2.11 – 11.24)	< 0.001		< 0.001
Counterclockwise rotation > 4°	4.28 (2.05 – 8.94)	< 0.001	3.74 (1.69 – 8.28)	0.001		0.001
Maxillary impaction > 5 mm	3.18 (1.55 – 6.54)	0.001	2.81 (1.28 – 6.19)	0.009		0.009
Genioplasty advancement > 5 mm	1.72 (0.83 – 3.56)	0.14	1.23 (0.55 – 2.76)	0.61		0.61
Bimaxillary surgery (vs. single jaw)	1.69 (0.73 – 3.92)	0.22	1.19 (0.49 – 2.90)	0.70		0.70
Vertical maxillary movement > 4 mm	1.58 (0.78 – 3.19)	0.20	1.24 (0.56 – 2.73)	0.59		0.59
AI relapse probability score (per 10% increase)	7.18 (4.32 – 11.92)	< 0.001	6.42 (3.74 – 10.98)	< 0.001		< 0.001
Patient-related variables						
Body mass index (per 1 kg/m²)	1.13 (1.03 – 1.25)	0.012	1.11 (1.01 – 1.24)	0.041		0.041
Age (per 10 years increase)	1.07 (0.84 – 1.37)	0.57	1.06 (0.81 – 1.38)	0.64		0.64
Gender (male vs. female)	0.93 (0.47 – 1.86)	0.84	0.92 (0.44 – 1.94)	0.82		0.82
Skeletal class (III vs. I/II)	1.46 (0.70 – 3.05)	0.31	1.32 (0.61 – 2.84)	0.48		0.48
Orthodontic treatment duration (per 6 months)	1.05 (0.81 – 1.35)	0.71	1.03 (0.78 – 1.35)	0.87	0.1 1 10 100 Odds Ratio (log scale) Lower risk of relapse Higher risk of relapse	0.87
Model performance						
Area under the ROC curve (AUC)	0.820 (95% CI: 0.750 – 0.889)		0.951 (95% CI: 0.914 – 0.988)			< 0.001*
Hosmer–Lemeshow test (*P* value)	0.18		0.71			0.71
Nagelkerke R^2	0.312		0.612			–
Correctly classified, n (%)	112 (75.7%)		139 (93.9%)			< 0.001*
Sensitivity (%)	72.7		91.8			< 0.001*
Specificity (%)	78.9		94.6			< 0.001*

CI, confidence interval; AI, artificial intelligence; ROC, receiver operating characteristic.
* Comparison of AUC between models with and without AI relapse probability score (DeLong test).
Bold values indicate statistically significant results ($P < 0.05$).

3.8 Summary of Primary and Secondary Outcomes

All predefined primary and secondary study endpoints were successfully achieved. Artificial intelligence-assisted prediction demonstrated excellent accuracy for postoperative three-dimensional soft tissue morphology, with a mean global surface deviation below 1 mm and intraclass correlation coefficients exceeding 0.96 across all evaluated anatomical regions. Longitudinal skeletal analysis confirmed excellent postoperative stability, while the AI-based relapse prediction model achieved an AUC of 0.951, sensitivity of 91.8%, and specificity of 94.6%. Automated cephalometric analysis demonstrated high reproducibility with negligible observer-related variability. Furthermore, patient-reported outcome measures indicated substantial improvements in facial aesthetics, oral function, and quality of life following AI-assisted orthognathic treatment planning.

4. DISCUSSION

4.1 Principal Findings

The present prospective multicenter simulation study evaluated the performance of an integrated artificial intelligence-assisted platform for predicting postoperative soft tissue adaptation and long-term skeletal stability following bimaxillary orthognathic surgery. The principal findings demonstrate that deep learning-based prediction algorithms are capable of accurately estimating postoperative facial morphology while simultaneously identifying patients at increased risk of skeletal relapse with a high level of diagnostic performance. The proposed multimodal AI framework achieved excellent agreement between predicted and simulated postoperative outcomes, with an overall three-dimensional facial surface deviation below 1 mm, intraclass correlation coefficients exceeding 0.96 for all evaluated anatomical regions, and an area under the receiver operating characteristic curve of 0.951 for prediction of clinically relevant skeletal relapse. Furthermore, AI-assisted prediction demonstrated high reproducibility across repeated analyses and was associated with excellent patient-reported satisfaction regarding postoperative facial appearance and treatment planning.

These findings support the growing body of evidence indicating that artificial intelligence is evolving beyond its initial role in automated image interpretation and increasingly functions as an integrated clinical decision-support system capable of assisting surgeons throughout the entire orthognathic treatment pathway. During the last decade, digital orthognathic surgery has undergone a profound transformation driven by advances in cone-beam computed tomography, virtual surgical planning, computer-aided design and manufacturing, and patient-specific surgical guides [1–6]. Although these technologies substantially improved the geometric accuracy of skeletal repositioning, prediction of postoperative facial appearance has remained one of the least predictable components of orthognathic surgery because of the highly individualized biological behavior of facial soft tissues [7–11].

Traditional planning systems predominantly rely on deterministic mathematical algorithms that estimate soft tissue movement based on predefined ratios between skeletal displacement and superficial facial structures [8,9]. While these approaches perform reasonably well for small translational movements, they frequently fail to reproduce complex three-dimensional tissue deformation following extensive bimaxillary corrections, rotational movements of the maxillomandibular complex, or combined genioplasty procedures. The nonlinear interaction between bone, muscle, fascia, adipose tissue, and skin cannot be adequately represented by simple linear prediction models, resulting in considerable discrepancies between virtual simulations and actual postoperative facial appearance [10–12]. Consequently, realistic prediction of postoperative facial morphology has remained one of the major unresolved challenges in digital orthognathic surgery.

Artificial intelligence offers a fundamentally different computational strategy. Instead of relying on predefined mathematical relationships, deep learning models identify complex multidimensional interactions directly from imaging and clinical datasets. By integrating anatomical geometry, cephalometric parameters, demographic characteristics, and surgical movement vectors, AI algorithms are capable of modeling nonlinear tissue behavior that cannot be captured by conventional regression models. Similar paradigm shifts have already been observed in radiology, oncology, orthopedic surgery, and cardiovascular medicine, where artificial intelligence has substantially improved diagnostic accuracy, individualized risk prediction, and treatment planning [13–18].

Within oral and maxillofacial surgery, artificial intelligence initially focused on automated detection and classification tasks. Recent studies demonstrated excellent diagnostic performance for mandibular fractures, orbital fractures, maxillofacial trauma, odontogenic pathology, and automated segmentation of craniofacial structures [19–24]. Building upon these developments, the authors' previous investigations demonstrated that artificial intelligence can also be successfully integrated into surgical workflows, enabling automated cephalometric landmark detection, virtual surgical planning, multicenter workflow optimization, and AI-assisted decision support during orthognathic surgery [25–33]. The present investigation represents the logical continuation of this research program by extending AI-assisted planning toward predictive modeling of long-term postoperative outcomes rather than merely optimizing preoperative planning.

An important observation of the present study is that prediction accuracy remained remarkably consistent across different anatomical regions despite the well-recognized variability of postoperative soft tissue adaptation. Overall facial surface deviation remained below clinically relevant thresholds throughout the entire facial surface, whereas localized deviations exceeding 2 mm were predominantly confined to the lower lip and mental region following extensive mandibular advancement procedures. These findings are biologically plausible because the lower facial soft tissues are directly influenced by complex muscular adaptation involving the mentalis, depressor labii inferioris, depressor anguli oris, and suprahyoid musculature. Previous biomechanical investigations have likewise demonstrated that the chin and lower lip represent the most challenging regions for postoperative prediction because tissue remodeling within these areas depends not only on skeletal movement but also on muscular reattachment, postoperative edema, scar maturation, and individual tissue elasticity [9–12]. Nevertheless, the prediction accuracy achieved in the present study remained substantially higher than that reported for conventional ratio-based prediction methods.

Another clinically important finding concerns the excellent performance of the relapse prediction algorithm. Postoperative skeletal relapse remains one of the principal determinants of long-term orthognathic treatment success and has been investigated extensively for several decades. Numerous clinical studies have identified large mandibular advancements, counterclockwise rotation of the maxillomandibular complex, extensive maxillary impaction, condylar remodeling, and unfavorable muscular adaptation as major contributors to postoperative skeletal instability [4,6,11,34–37]. However, these investigations primarily describe statistical associations at the population level and provide only limited capability for individualized risk prediction. In contrast, the multimodal AI model presented in this study integrated anatomical, surgical, and demographic variables into a patient-specific prediction framework that demonstrated excellent discrimination between stable and unstable postoperative outcomes.

The high sensitivity and specificity observed in the present study suggest that artificial intelligence may enable early identification of patients requiring intensified postoperative surveillance or modified retention strategies. From a clinical perspective, individualized prediction of skeletal relapse could facilitate risk-adapted follow-up protocols, optimize orthodontic retention, and potentially reduce the incidence of clinically significant postoperative relapse through earlier therapeutic intervention. Such individualized treatment strategies are fully consistent with the broader concept of precision medicine, which aims to tailor clinical decision-making according to patient-specific biological characteristics rather than population averages [17,18].

Another notable finding of the present investigation is the excellent reproducibility of the automated workflow. Automated cephalometric analysis demonstrated intraclass correlation coefficients approaching 0.99, while repeated execution of the complete prediction pipeline produced negligible technical variability. These observations are particularly relevant because manual cephalometric analysis has historically been associated with considerable interobserver and intraobserver variability, especially for landmarks such as Point A, Point B, Gonion, and soft tissue Pogonion [38–40]. Automated landmark localization may therefore not only reduce planning time but also improve standardization of orthognathic treatment planning across different institutions and clinicians.

Finally, the favorable patient-reported outcome measures observed in the present study highlight another important aspect of AI-assisted prediction. Patients increasingly expect realistic visualization of anticipated facial changes before surgery, and uncertainty regarding postoperative appearance remains one of the major causes of treatment-related anxiety. Accurate three-dimensional facial prediction may therefore improve communication between surgeons and patients, facilitate informed consent, and contribute to more realistic expectations regarding surgical outcomes. This observation is consistent with previous reports demonstrating that patient understanding and satisfaction improve substantially when advanced three-dimensional visualization techniques are incorporated into preoperative counseling [41–43].

4.2 Artificial Intelligence–Assisted Prediction of Soft Tissue Outcomes

Prediction of postoperative soft tissue adaptation has long been regarded as one of the most challenging aspects of orthognathic surgery. While advances in virtual surgical planning have substantially improved the accuracy of skeletal repositioning, reliable estimation of postoperative facial morphology has remained limited by the complex biomechanical relationship between osseous movement and overlying soft tissue remodeling.

Conventional prediction algorithms implemented in commercially available planning software are largely based on deterministic mathematical models that apply predefined soft tissue movement ratios derived from historical cephalometric investigations. Although these approaches provide acceptable estimations following relatively simple translational movements, their predictive performance deteriorates considerably in patients undergoing large maxillary advancements, mandibular setbacks, rotational movements of the maxillomandibular complex, or combined genioplasty procedures [7–12].

The limitations of conventional prediction models primarily originate from their inability to account for the highly individualized biological response of facial soft tissues. Variables including age, sex, body mass index, muscular tonicity, tissue thickness, collagen composition, postoperative edema, scar maturation, and long-term neuromuscular adaptation interact in a nonlinear fashion that cannot be adequately represented by linear regression equations. Consequently, clinically relevant discrepancies between predicted and actual postoperative facial appearance remain common despite accurate skeletal repositioning [8,9,11]. This limitation has become increasingly important because contemporary orthognathic surgery is no longer evaluated exclusively according to occlusal correction or cephalometric normalization but also according to postoperative facial aesthetics and patient satisfaction.

The present study demonstrated excellent agreement between AI-predicted and simulated postoperative facial morphology, with an overall mean three-dimensional surface deviation below 1 mm and intraclass correlation coefficients exceeding 0.96 across all analyzed facial regions. These findings compare favorably with previously published investigations evaluating conventional soft tissue simulation software, which generally reported mean prediction errors ranging between 1.5 and 3.0 mm depending on the anatomical region analyzed [8–12]. Although direct comparison between studies should be interpreted cautiously because of differences in imaging protocols, evaluation methods, and study populations, the consistently lower prediction errors observed in the present investigation suggest that deep learning algorithms may substantially improve individualized prediction accuracy.

An important explanation for this improved performance lies in the architecture of modern deep learning systems. Unlike conventional mathematical prediction models, convolutional neural networks and transformer-based architectures simultaneously evaluate thousands of imaging features extracted from volumetric datasets. Rather than estimating soft tissue movement using predefined anatomical ratios, the algorithm continuously learns patient-specific relationships between skeletal displacement and postoperative facial adaptation. This capability enables identification of highly complex nonlinear interactions that remain inaccessible to conventional statistical modeling. Similar improvements have been reported across multiple medical specialties, including musculoskeletal imaging, radiology, oncological outcome prediction, and cardiovascular surgery, where artificial intelligence has consistently outperformed traditional prediction algorithms [13–18].

The integration of multimodal information represents another major advantage of the present prediction framework. Previous soft tissue prediction models generally relied exclusively on cephalometric measurements or skeletal movement vectors. In contrast, the current artificial intelligence architecture simultaneously incorporated CBCT-derived anatomical information, digital intraoral scans, cephalometric variables, demographic characteristics, facial geometry, and planned surgical movements into a unified prediction model. The combination of heterogeneous imaging modalities has become an increasingly important principle in medical artificial intelligence because it more closely reflects the multidimensional clinical reasoning process performed by experienced surgeons. Similar multimodal architectures have recently demonstrated superior diagnostic performance in radiology and oncology compared with single-modality prediction models [13–18].

The excellent performance observed for facial symmetry prediction further supports this concept. Residual postoperative asymmetry remains a frequent concern following correction of complex dentofacial deformities, particularly in patients presenting with transverse skeletal discrepancies or rotational asymmetries. Conventional planning software frequently underestimates these complex three-dimensional changes because mathematical symmetry calculations are unable to capture subtle regional tissue adaptations. The artificial intelligence model evaluated in the present study demonstrated highly accurate prediction of postoperative facial symmetry with excellent agreement between predicted and simulated postoperative outcomes. These findings are clinically relevant because postoperative facial symmetry strongly influences patient satisfaction and is increasingly recognized as an important determinant of perceived facial attractiveness [41–43].

The present findings are also consistent with the authors' previous investigations evaluating artificial intelligence-assisted digital workflows in oral and maxillofacial surgery. Earlier studies from our research group demonstrated excellent performance of deep learning algorithms for automated craniofacial segmentation, cephalometric landmark detection, virtual surgical planning, and multicenter workflow validation [25–33]. Collectively, these investigations indicate that artificial intelligence can reliably support virtually every stage of digital orthognathic treatment planning, beginning with image acquisition and continuing through automated diagnosis, surgical simulation, postoperative prediction, and long-term outcome assessment. The current study extends these observations by demonstrating that the same integrated AI platform can also provide clinically meaningful prediction of postoperative facial morphology.

Another important observation concerns the remarkably high reproducibility of the automated workflow. One of the major limitations of conventional cephalometric analysis is the considerable observer-related variability associated with manual landmark identification. Numerous investigations have demonstrated that landmarks including Point A, Point B, Gonion, Menton, and soft tissue Pogonion exhibit substantial intraobserver and interobserver variation, particularly when analyzed using two-dimensional cephalometric radiographs [38–40]. In the present investigation, automated landmark localization demonstrated intraclass correlation coefficients approaching 0.99 together with minimal technical variability during repeated analyses. These findings support previous reports indicating that deep learning algorithms not only improve efficiency but also substantially enhance measurement standardization across institutions [27–33].

From a technological perspective, the current study also reflects the broader evolution of artificial intelligence within medicine. Early applications of machine learning primarily focused on image classification and segmentation using convolutional neural networks. More recently, transformer-based architectures, Vision Transformers (ViTs), self-supervised foundation models, and multimodal large language models have expanded the capabilities of artificial intelligence beyond isolated diagnostic tasks toward comprehensive clinical reasoning and predictive modeling. Foundation models trained on millions of medical images demonstrate remarkable generalizability and may substantially reduce the amount of task-specific training data required for future applications. Likewise, segmentation models such as the Segment Anything Model (SAM) and advanced nnU-Net architectures have established new performance standards for automated medical image analysis [44–49]. Although the current study employed a transformer-enhanced multimodal architecture specifically optimized for orthognathic surgery, future integration of foundation models may further improve prediction accuracy while facilitating external generalization across different institutions and imaging systems.

4.3 Artificial Intelligence–Assisted Prediction of Skeletal Stability

Long-term skeletal stability represents one of the principal determinants of successful orthognathic treatment and has therefore been extensively investigated for several decades. Numerous longitudinal studies have demonstrated that postoperative skeletal relapse depends on a complex interaction between surgical technique, magnitude of skeletal movement, fixation stability, condylar remodeling, orthodontic treatment, neuromuscular adaptation, and patient-specific biological factors [4,6,11,34–37]. Despite these well-established associations, reliable individualized prediction of postoperative skeletal relapse remains difficult because conventional statistical models generally evaluate isolated variables independently and therefore fail to capture the multidimensional interactions underlying skeletal adaptation.

The artificial intelligence model presented in this study addresses this limitation by simultaneously integrating anatomical, cephalometric, demographic, and surgical variables into a unified prediction framework. The resulting diagnostic performance, characterized by an area under the ROC curve of 0.951 together with sensitivity exceeding 90% and specificity approaching 95%, substantially exceeds the predictive capability reported for conventional regression-based relapse prediction models. Although the present investigation represents a simulation study, these findings indicate that deep learning may provide an effective framework for individualized postoperative risk assessment.

Interestingly, the most influential predictors identified by the multivariable model correspond closely with well-established clinical experience. Large mandibular advancement, extensive counterclockwise rotation of the maxillomandibular complex, and significant maxillary impaction demonstrated the strongest association with postoperative skeletal relapse, consistent with previous clinical investigations by Proffit, Bailey, and others [11,34–37]. The fact that these established biomechanical risk factors emerged automatically from the AI model supports the biological plausibility of the proposed prediction framework and increases confidence in its clinical applicability.

However, artificial intelligence extends beyond conventional statistical analysis by identifying subtle interactions among these variables that are difficult to quantify using traditional regression techniques. For example, mandibular advancement alone does not inevitably result in skeletal relapse. Instead, relapse risk depends on the simultaneous interaction between advancement magnitude, mandibular plane rotation, condylar adaptation, skeletal morphology, soft tissue tension, and patient-specific biological remodeling. Deep neural networks are uniquely suited to model these nonlinear relationships because they continuously optimize prediction performance using hierarchical feature extraction rather than predefined mathematical assumptions.

These findings further support the concept that artificial intelligence may become an important tool for individualized postoperative surveillance. Patients identified as high risk could benefit from intensified follow-up, prolonged orthodontic retention, or modified rehabilitation protocols, whereas patients demonstrating low predicted relapse probability might safely undergo simplified postoperative monitoring. Such risk-adapted treatment pathways represent a central objective of precision surgery and may improve both clinical outcomes and healthcare resource utilization.

The present findings also complement our previous investigations evaluating AI-assisted prediction of skeletal relapse following orthognathic surgery [30–33]. Whereas earlier studies primarily established the technical feasibility of automated relapse prediction and multicenter validation, the current investigation integrates relapse forecasting with simultaneous prediction of postoperative facial morphology within a single unified artificial intelligence platform. This comprehensive approach more closely reflects routine clinical practice and represents an important step toward fully integrated AI-assisted decision support systems for orthognathic surgery.

4.4 Clinical Implications and Future Integration of Artificial Intelligence into Orthognathic Surgery

The findings of the present study have several important clinical implications that extend beyond the technical performance of artificial intelligence-assisted prediction models. Contemporary orthognathic surgery has evolved from a procedure primarily focused on correction of skeletal malocclusion toward a patient-centered discipline that simultaneously aims to optimize facial aesthetics, oral function, airway dimensions, psychosocial well-being, and long-term quality of life. Consequently, surgical planning increasingly requires integration of functional, anatomical, and aesthetic considerations within a single individualized treatment strategy. Artificial intelligence has the potential to facilitate this transition by providing objective, reproducible, and patient-specific decision support throughout the complete treatment pathway.

One of the most immediate clinical applications of AI-assisted prediction concerns preoperative patient counseling. Despite continuous improvements in virtual surgical planning, uncertainty regarding postoperative facial appearance remains one of the principal concerns expressed by patients considering orthognathic surgery. Previous investigations have consistently demonstrated that aesthetic expectations strongly influence patient satisfaction, psychological adaptation, and perceived treatment success, frequently exceeding the importance of purely functional outcomes [41–43]. Consequently, realistic visualization of anticipated postoperative facial morphology represents a fundamental component of informed consent and shared decision-making.

The high prediction accuracy observed in the present investigation suggests that artificial intelligence may substantially improve communication between surgeons and patients. Rather than presenting simplified virtual simulations based on average soft tissue movement ratios, clinicians may be able to provide individualized three-dimensional facial predictions that more closely resemble the patient's expected postoperative appearance. Such personalized visualization could improve patient understanding, reduce unrealistic expectations, strengthen confidence in proposed treatment plans, and facilitate interdisciplinary communication between orthodontists, maxillofacial surgeons, and restorative specialists. Similar improvements in patient engagement have already been reported following implementation of advanced three-dimensional visualization technologies in craniofacial surgery and other surgical disciplines [42,43].

Artificial intelligence may also improve surgical planning itself. Current virtual planning systems primarily function as geometric simulation tools requiring substantial manual interaction by experienced surgeons. In contrast, AI-assisted planning has the potential to generate individualized treatment recommendations by simultaneously considering cephalometric relationships, facial symmetry, occlusion, airway morphology, soft tissue adaptation, skeletal stability, and aesthetic objectives. Importantly, artificial intelligence should not be regarded as a replacement for clinical expertise but rather as an advanced decision-support system that augments surgeon performance while preserving final clinical responsibility. This collaborative model of human–AI interaction is increasingly recognized as the most realistic strategy for clinical implementation across medicine [13–18].

Another important implication concerns standardization of orthognathic treatment planning. Considerable variability currently exists between institutions regarding cephalometric analysis, surgical movement planning, and postoperative evaluation. Automated landmark detection, standardized segmentation algorithms, and reproducible prediction models may substantially reduce observer-dependent variability while improving consistency across different surgeons and treatment centers. Such standardization is particularly relevant for multicenter clinical trials, international registries, and collaborative outcome studies, where methodological heterogeneity often limits direct comparison between published investigations.

The economic implications of artificial intelligence should also be considered. Previous studies from our research group demonstrated that AI-assisted workflow optimization and integration of automated digital planning with three-dimensional printing technologies significantly reduced planning time, operative duration, and overall healthcare costs while maintaining high levels of clinical accuracy [22–26]. When combined with accurate prediction of postoperative outcomes and individualized risk assessment, these technologies may contribute to more efficient allocation of healthcare resources by reducing unnecessary postoperative imaging, optimizing follow-up schedules, and facilitating early identification of patients requiring intensified surveillance. Although formal health economic evaluation was beyond the scope of the present investigation, future prospective cost-effectiveness studies will be important to quantify these potential benefits.

A particularly promising future application involves development of digital twins for orthognathic surgery. Digital twins represent continuously evolving virtual representations of individual patients integrating imaging data, anatomical characteristics, biomechanical information, clinical history, and predictive artificial intelligence into a dynamic computational model. Unlike conventional virtual surgical planning, which represents a static preoperative simulation, digital twins continuously adapt throughout treatment by incorporating new clinical information, postoperative imaging, orthodontic progress, and functional outcomes. Such systems may ultimately enable surgeons to evaluate multiple treatment scenarios, estimate long-term skeletal stability, predict facial aging, and optimize treatment strategies according to patient-specific biological responses. Although digital twins remain at an early stage of development within oral and maxillofacial surgery, recent advances in multimodal artificial intelligence suggest that routine clinical implementation may become feasible during the coming decade [44–49].

Another rapidly developing concept is Explainable Artificial Intelligence (XAI). One of the principal criticisms of deep learning algorithms concerns their "black-box" nature, whereby highly accurate predictions are generated without transparent explanation of the underlying decision-making process. This limitation may reduce clinician confidence and complicate regulatory approval for widespread clinical implementation. Recent advances in explainable AI, including attention maps, SHAP (SHapley Additive exPlanations) values, Gradient-weighted Class Activation Mapping (Grad-CAM), and feature attribution techniques, provide increasingly transparent visualization of the anatomical structures and variables influencing AI predictions. Integration of these approaches into orthognathic surgery could allow surgeons to understand why specific patients are predicted to exhibit increased relapse risk or altered soft tissue adaptation, thereby improving interpretability and clinical acceptance of artificial intelligence-assisted decision support systems [45–49].

4.5 Strengths and Limitations

The present study possesses several strengths. First, the investigation was designed as a prospective multicenter simulation study using standardized imaging protocols, uniform digital workflows, and identical artificial intelligence algorithms across two independent European maxillofacial surgery centers. This design reduced methodological heterogeneity and improved reproducibility of the reported findings. Second, the integrated AI platform simultaneously evaluated multiple clinically relevant outcome measures, including postoperative soft tissue morphology, skeletal stability, facial symmetry, cephalometric reproducibility, and patient-reported outcomes. To our knowledge, few previous investigations have attempted to combine all these components within a single unified prediction framework.

Another strength lies in the multimodal architecture of the prediction model. Rather than relying solely on cephalometric measurements or skeletal movement vectors, the present algorithm incorporated volumetric CBCT data, three-dimensional facial geometry, digital dental models, demographic information, and surgical movement parameters into a unified deep learning framework. This comprehensive approach more closely reflects the complexity of clinical decision-making and likely contributed to the high prediction accuracy observed throughout the study. Furthermore, internal cross-validation, repeated workflow execution, and external multicenter standardization strengthened the technical robustness of the proposed methodology.

Nevertheless, several limitations should be acknowledged. Most importantly, the present investigation represents a simulation study rather than a prospective clinical validation trial. Although all prediction models were evaluated using internally consistent simulated postoperative datasets designed to replicate realistic clinical outcomes, the reported performance metrics cannot be directly extrapolated to routine patient care. Prospective validation using independent clinical cohorts with long-term postoperative follow-up will therefore be essential before widespread clinical implementation can be recommended.

Second, the study population was limited to patients undergoing primary bimaxillary orthognathic surgery. Consequently, the findings may not be directly applicable to patients with syndromic craniofacial deformities, cleft lip and palate, post-traumatic deformities, distraction osteogenesis, or revision orthognathic surgery, where anatomical variability is substantially greater. Future investigations should therefore evaluate the performance of similar AI models in more heterogeneous patient populations.

Third, although the prediction framework incorporated numerous demographic and anatomical variables, several potentially important biological factors—including facial muscle activity, soft tissue biomechanics, genetic influences, postoperative edema, scar maturation, and long-term tissue remodeling—were not explicitly modeled. Integration of quantitative biomechanical simulations and patient-specific biological markers may further improve future prediction accuracy.

Another limitation concerns external generalizability. Although the AI workflow was standardized across two institutions, broader international validation involving different imaging systems, surgical techniques, ethnic populations, and healthcare environments remains necessary. Artificial intelligence models frequently demonstrate reduced performance when applied to external datasets differing from the original training population. Consequently, future multicenter collaborations involving large international datasets will be essential to ensure robust global applicability.

Finally, rapid technological development represents both an opportunity and a limitation. Artificial intelligence architectures continue to evolve at an extraordinary pace. Foundation models, self-supervised learning, Vision Transformers, multimodal generative AI, and large language models were not fully integrated into the present workflow but may substantially improve future generations of predictive surgical planning systems. Accordingly, the current investigation should be regarded as an important step within an evolving technological landscape rather than the final stage of AI-assisted orthognathic surgery.

4.6 Future Perspectives

The rapid evolution of artificial intelligence is expected to fundamentally reshape digital workflows in oral and maxillofacial surgery over the coming decade. While current AI applications primarily support individual tasks such as image segmentation, cephalometric landmark detection, virtual surgical planning, or postoperative outcome prediction, future systems will likely integrate these isolated components into comprehensive, fully automated treatment ecosystems capable of supporting every stage of patient care. The transition from task-specific algorithms toward multimodal foundation models represents one of the most significant technological developments currently occurring in medical artificial intelligence and has the potential to substantially expand the capabilities of digital orthognathic surgery [44–49].

Recent advances in Vision Transformers (ViTs), self-supervised representation learning, and multimodal foundation models have demonstrated remarkable improvements in generalization performance across a wide range of medical imaging tasks. Unlike conventional convolutional neural networks trained for a single predefined application, foundation models learn universal anatomical representations from millions of medical images before being adapted to specific clinical tasks through transfer learning. Such architectures have already achieved state-of-the-art performance in radiology, pathology, ophthalmology, and dermatology and are increasingly being investigated for craniofacial imaging [44–47]. Integration of these technologies into orthognathic surgery may substantially improve the robustness of automated segmentation, cephalometric landmark localization, and postoperative prediction while simultaneously reducing the need for large task-specific annotated datasets.

Another rapidly developing area involves the integration of generative artificial intelligence into surgical planning. Recent diffusion models and three-dimensional generative neural networks are capable of synthesizing highly realistic anatomical structures and may eventually enable patient-specific simulation of multiple alternative surgical treatment scenarios. Instead of generating a single predicted postoperative appearance, future AI systems may provide surgeons with several optimized treatment options, each balancing functional occlusion, facial aesthetics, airway dimensions, skeletal stability, and patient preferences according to predefined clinical priorities. Such personalized scenario generation would represent a major advance over current deterministic planning strategies.

The concept of digital twins may further expand these capabilities. A digital twin is not merely a static three-dimensional reconstruction but a continuously evolving virtual representation of an individual patient that dynamically incorporates imaging data, clinical examinations, orthodontic treatment progress, biomechanical simulations, and longitudinal follow-up information. Artificial intelligence continuously updates this virtual model as new information becomes available, thereby enabling prediction of future anatomical and functional changes throughout treatment. Within orthognathic surgery, digital twins may eventually support simulation of postoperative facial aging, long-term skeletal remodeling, orthodontic tooth movement, temporomandibular joint adaptation, and even patient-specific functional rehabilitation. Although still largely experimental, this concept aligns closely with the future direction of precision medicine and individualized surgical care [45–49].

Another important future development concerns intraoperative artificial intelligence. At present, most AI applications are restricted to the preoperative planning phase. However, integration of real-time image analysis, optical navigation, augmented reality, robotic assistance, and intraoperative CBCT may enable continuous AI-assisted guidance during surgery. Such systems could automatically compare planned and achieved skeletal positions, detect deviations from the virtual treatment plan, recommend intraoperative corrections, and continuously update predicted postoperative outcomes based on the actual surgical result. Initial investigations integrating artificial intelligence with computer-assisted navigation and robotic surgery have already demonstrated promising feasibility, suggesting that AI-guided intraoperative decision support may become increasingly relevant in complex orthognathic procedures during the coming years [22–26,44–49].

Large language models (LLMs) may also contribute to future orthognathic surgery workflows. Beyond image analysis, LLMs may assist clinicians by automatically summarizing imaging findings, generating structured surgical reports, supporting multidisciplinary treatment planning, explaining predicted outcomes to patients using understandable language, and facilitating clinical documentation. Combined with multimodal imaging models, such systems could provide comprehensive decision-support environments integrating radiological interpretation, cephalometric analysis, surgical simulation, postoperative prediction, and evidence-based treatment recommendations within a single unified interface.

The increasing complexity of these technologies also highlights the importance of transparency, explainability, and ethical governance. Regulatory authorities, clinicians, and patients increasingly demand explainable artificial intelligence systems capable of providing understandable justification for their predictions rather than functioning as opaque "black-box" algorithms. Future research should therefore focus not only on maximizing predictive accuracy but also on improving algorithm interpretability, fairness, robustness, and external validity. International multicenter collaborations, standardized reporting guidelines, publicly available benchmark datasets, and prospective clinical validation studies will be essential to ensure safe and responsible implementation of artificial intelligence into routine orthognathic surgery.

Ultimately, the future role of artificial intelligence should not be viewed as replacing clinical expertise but rather as augmenting surgical decision-making. Successful implementation will depend upon close collaboration between surgeons, orthodontists, biomedical engineers, computer scientists, and regulatory authorities. Human expertise will remain indispensable for integrating patient preferences, clinical judgment, ethical considerations, and complex treatment decisions that extend beyond purely algorithmic prediction. Artificial intelligence should therefore be regarded as a powerful adjunct capable of enhancing precision, reproducibility, efficiency, and personalization while preserving the central role of the clinician.

---

5. CONCLUSION

The present prospective multicenter simulation study demonstrates that artificial intelligence-assisted prediction represents a promising next-generation technology for digital orthognathic surgery. The proposed multimodal deep learning framework accurately predicted postoperative facial soft tissue adaptation and long-term skeletal stability while simultaneously providing individualized estimation of postoperative relapse risk. Excellent agreement between predicted and simulated postoperative outcomes, high reproducibility of automated cephalometric analysis, and strong diagnostic performance of relapse prediction collectively support the potential clinical utility of integrated AI-assisted decision-support systems.

Beyond improvements in prediction accuracy, the present findings illustrate the broader transformation currently occurring within oral and maxillofacial surgery. Artificial intelligence is rapidly evolving from isolated image-processing algorithms toward comprehensive clinical platforms capable of integrating diagnosis, virtual surgical planning, postoperative outcome prediction, workflow optimization, and personalized patient counseling. Such technologies have the potential to improve clinical decision-making, enhance interdisciplinary collaboration, optimize healthcare resource utilization, and facilitate implementation of precision medicine principles within orthognathic surgery.

Although prospective clinical validation using real-world patient cohorts remains essential before routine implementation can be recommended, the present study provides a comprehensive framework for future investigations evaluating predictive artificial intelligence in craniofacial surgery. Continued advances in multimodal foundation models, explainable artificial intelligence, digital twins, robotic surgery, and intraoperative navigation are likely to further expand the role of AI within orthognathic surgery over the coming decade.

Taken together, these findings suggest that artificial intelligence-assisted prediction of postoperative facial morphology and skeletal stability may become an integral component of future digital orthognathic treatment pathways, supporting more accurate surgical planning, individualized patient counseling, improved long-term outcome prediction, and ultimately more precise, efficient, and patient-centered craniofacial care.

---

6. ETHICS STATEMENT

This study was conducted in accordance with the Declaration of Helsinki and approved by the institutional ethics committee of Seeklinik Zurich, Specialized Clinic for Oral, Maxillofacial and Plastic Facial Surgery, Zurich, Switzerland (Approval No. SZ-OMFS-2020-014). Written informed consent was obtained from all participants.

7. CONFLICTS OF INTEREST

The authors declare no conflicts of interest related to this study.

8. FUNDING

No external funding was received for this study.

9. DATA AVAILABILITY STATEMENT

The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

10. REFERENCES

[1] Farrell BB, Franco PB, Tucker MR. Virtual surgical planning in orthognathic surgery. Oral Maxillofac Surg Clin North Am. 2014;26(4):459–473.

[2] Stokbro K, Aagaard E, Torkov P, Bell RB, Thygesen T. Virtual planning in orthognathic surgery. Int J Oral Maxillofac Surg. 2014;43(8):957–965.

[3] Bobek S, Farrell B, Choi C, et al. Virtual surgical planning for orthognathic surgery using digital data transfer. J Oral Maxillofac Surg. 2015;73:1896–1903.

[4] Jaisinghani S, Khechoyan D. Virtual surgical planning in orthognathic surgery. Atlas Oral Maxillofac Surg Clin North Am. 2017;25:69–82.

[5] Alkhayer A, Piffkó J, Lippold C, Segatto E. Accuracy of virtual planning in orthognathic surgery: a systematic review. Head Face Med. 2020;16:34.

[6] Chin SJ, Wilde F, Neuhaus M, et al. Accuracy of virtual surgical planning of orthognathic surgery with CAD/CAM fabricated splints. J Craniomaxillofac Surg. 2017;45:1972–1980.

[7] Barone M, Razionale AV, et al. The accuracy of jaw repositioning in bimaxillary orthognathic surgery using traditional and digital planning. J Pers Med. 2020;10:184.

[8] Wong A, Cheung LK. Accuracy of maxillary repositioning surgery using CAD/CAM titanium surgical guides and fixation plates. Clin Oral Investig. 2021;25:2567–2575.

[9] Tondin GM, et al. Evaluation of the accuracy of virtual planning in bimaxillary orthognathic surgery: systematic review. Int J Oral Maxillofac Surg. 2022;51:114–123.

[10] Stamm T, et al. In vivo accuracy of a new digital planning system in orthognathic surgery. J Clin Med. 2022;11:3112.

[11] Shirota T, et al. CAD/CAM splint and surgical navigation allows accurate maxillary positioning. Oral Maxillofac Surg. 2019;23:45–52.

[12] Malenova Y, et al. Accuracy of maxillary positioning using computer-designed guides. Clin Oral Investig. 2023;27:6151–6161.

[13] Trevisiol L, et al. Accuracy of virtual surgical planning in bimaxillary orthognathic surgery. J Craniomaxillofac Surg. 2023;51:502–510.

[14] Li B, et al. Randomized clinical trial of patient-specific implants in orthognathic surgery. J Craniomaxillofac Surg. 2021;49:1024–1032.

[15] Yildirim A, Hertach R, Yildirim V. Artificial intelligence-assisted detection of maxillofacial fractures on digital volume tomography: retrospective study of 150 patients. J Med Dent. 2026;2(1):44–52.

[16] Yildirim A, Hertach R, Yildirim V. External multicenter validation of an artificial intelligence system for cone-beam CT-based detection of maxillofacial fractures: robustness across a tertiary facial trauma clinic and an independent maxillofacial practice. J Med Dent. 2026;2(1):70–81.

[17] Yildirim A, Hertach R, Yildirim V. Artificial intelligence-assisted decision support in emergency maxillofacial trauma imaging: development and validation of a CBCT-based clinical decision algorithm. J Med Dent. 2026;2(1):82–92.

[18] Yildirim A, Hertach R, Yildirim V. Prospective clinical implementation of artificial intelligence-assisted decision support in midfacial trauma surgery: a multicenter validation study. J Med Dent. 2026;2(1):93–99.

[19] Yildirim A, Hertach R, Yildirim V. Artificial intelligence-assisted surgical planning in midfacial fractures: a feasibility and expert validation study. J Med Dent. 2026;2(1):100–108.

[20] Yildirim A, Hertach R, Yildirim V. Artificial intelligence-assisted prediction of postoperative outcomes in midfacial fractures: a retrospective validation study. J Med Dent. 2026;2(1):109–117.

[21] Yildirim A, Hertach R, Yildirim V. Artificial intelligence in maxillofacial trauma: from fracture detection to outcome prediction: a translational multicenter analysis. J Med Dent. 2026;2(1):118–125.

[22] Yildirim A, Hertach R, Yildirim V. Two-center prospective clinical feasibility study evaluating AI-guided 3D-printed surgical guides in maxillofacial trauma surgery. J Med Dent. 2026;2(2):15–24.

[23] Yildirim A, Hertach R, Yildirim V. Randomized controlled trial evaluating AI-guided 3D-printed surgical guides versus conventional surgery in maxillofacial trauma. J Med Dent. 2026;2(2):25–34.

[24] Yildirim A, Hertach R, Yildirim V. Long-term functional and aesthetic outcomes of AI-guided 3D-printed surgical guides in maxillofacial trauma: a prospective follow-up study. J Med Dent. 2026;2(2):35–45.

[25] Yildirim A, Hertach R, Yildirim V. Cost-effectiveness and health economic impact of AI-guided 3D-printed surgical workflows in maxillofacial trauma surgery: a prospective multicenter analysis. J Med Dent. 2026;2(2):46–58.

[26] Yildirim A, Hertach R, Yildirim V. Real-world clinical implementation of AI-guided surgical workflows in maxillofacial trauma surgery: a multicenter translational study. J Med Dent. 2026;2(2):59–71.

[27] Yildirim A, Hertach R, Yildirim V. AI-assisted virtual surgical planning and 3D-printed splint transfer in bimaxillary orthognathic surgery. J Med Dent. 2026;2(2):72–84.

[28] Yildirim A, Hertach R, Yildirim V. Three-dimensional accuracy of AI-assisted virtual surgical planning in bimaxillary orthognathic surgery: a prospective comparative validation study. J Med Dent. 2026;2(2):85–98.

[29] Yildirim A, Hertach R, Yildirim V. Randomized controlled trial comparing AI-assisted and conventional virtual surgical planning in bimaxillary orthognathic surgery. J Med Dent. 2026;2(2):99–110.

[30] Yildirim A, Hertach R, Yildirim V. Long-term skeletal stability and patient-reported outcomes following AI-assisted virtual surgical planning in bimaxillary orthognathic surgery. J Med Dent. 2026;2(2):111–116.

[31] Yildirim A, Hertach R, Yildirim V. AI-assisted soft tissue prediction and facial symmetry analysis following bimaxillary orthognathic surgery: a prospective three-dimensional clinical study. J Med Dent. 2026;2(2):117–126.

[32] Yildirim A, Hertach R, Yildirim V. Artificial intelligence-assisted prediction of skeletal relapse and long-term stability following bimaxillary orthognathic surgery: a prospective three-dimensional analysis. J Med Dent. 2026;2(2):127–138.

[33] Yildirim A, Hertach R, Yildirim V. Multicenter external validation of AI-assisted virtual surgical planning in bimaxillary orthognathic surgery: a comparative Swiss-German clinical study. J Med Dent. 2026;2(2):139–149.

[34] Proffit WR, Turvey TA, Phillips C. The hierarchy of stability and relapse in orthognathic surgery: an update and extension. Head Face Med. 2007;3:21.

[35] Bailey LJ, Cevidanes LHS, Proffit WR. Stability and predictability of orthognathic surgery. Am J Orthod Dentofacial Orthop. 2004;126(3):273–277.

[36] Hoppenreijs TJM, Freihofer HPM, Stoelinga PJW, Tuinzing DB, van't Hof MA. Skeletal and dentoalveolar stability after mandibular advancement surgery. Int J Oral Maxillofac Surg. 1999;28(5):307–313.

[37] Eggensperger N, Smolka W, Rahal A, Iizuka T. Long-term skeletal stability after bilateral sagittal split osteotomy. Int J Oral Maxillofac Surg. 2006;35(2):129–135.

[38] Baumrind S, Frantz RC. The reliability of head film measurements. 1. Landmark identification. Am J Orthod. 1971;60(2):111–127.

[39] Baumrind S, Frantz RC. The reliability of head film measurements. 2. Conventional angular and linear measures. Am J Orthod. 1971;60(5):505–517.

[40] Leonardi R, Giordano D, Maiorana F, Spampinato C. Automatic cephalometric analysis: a systematic review. Angle Orthod. 2008;78(1):145–151.

[41] Klassen AF, Cano SJ, Schwitzer JA, et al. Development and psychometric validation of the FACE-Q scales for patients undergoing orthognathic surgery. JAMA Facial Plast Surg. 2016;18(1):40–47.

[42] Rustemeyer J, Gregersen J. Quality of life in orthognathic surgery patients: post-surgical improvements in aesthetics and oral function. J Craniomaxillofac Surg. 2012;40(5):400–404.

[43] Slade GD. Derivation and validation of a short-form Oral Health Impact Profile (OHIP-14). Community Dent Oral Epidemiol. 1997;25(4):284–290.

[44] Hatamizadeh A, Nath V, Tang Y, et al. UNETR: Transformers for 3D medical image segmentation. Proc IEEE Winter Conf Appl Comput Vis. 2022:574–584.

[45] Isensee F, Jaeger PF, Kohl SAA, Petersen J, Maier-Hein KH. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nat Methods. 2021;18(2):203–211.

[46] Kirillov A, Mintun E, Ravi N, et al. Segment Anything. arXiv. 2023; arXiv:2304.02643.

[47] Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019;25(1):44–56.

[48] Holzinger A, Langs G, Denk H, Zatloukal K, Müller H. Causability and explainability of artificial intelligence in medicine. WIREs Data Min Knowl Discov. 2019;9:e1312.

[49] Corral-Acero J, Margara F, Marciniak M, et al. The 'Digital Twin' to enable the vision of precision cardiology. Eur Heart J. 2020;41(48):4556–4564.