EFFECTIVENESS OF ARTIFICIAL INTELLIGENCE IN PREDICTING SOFT TISSUE OUTCOMES IN PATIENTS UNDERGOING ORTHOGNATHIC SURGERY: A SYSTEMATIC REVIEW

Chonticha CHOOKIARTSIRI[1], Chidchanok SESSIRISOMBAT[1] and Irin SIRISOONTORN[1]
1 Walailak University International College of Dentistry, Thailand;
chonticha.chookiartsir@gmail.com (C. C.); dr.nokorthodontist@gmail.com (C. S.);
irin.sirisoontorn@gmail.com (I. S.)

ARTICLE HISTORY
Received: 9 June 2025 **Revised:** 23 June 2025 **Published:** 7 July 2025

ABSTRACT
Predicting soft tissue changes after orthognathic surgery is essential but challenging, as traditional methods often lack accuracy in capturing individual patient outcomes. Recent advancements in artificial intelligence (AI) offer promising solutions for achieving more precise predictions. This systematic review examines the effectiveness of AI models in forecasting postoperative soft tissue changes in patients undergoing orthognathic surgery. A comprehensive literature search was performed in PubMed, Scopus, and ScienceDirect, identifying 22 relevant studies published from 2014 to April 2025 that compared AI predictions with actual postoperative results. These studies primarily utilized AI methods such as deep learning, finite element modeling (FEM), and landmark-based techniques, achieving prediction accuracy within errors ranging from 0.55 mm to 2.91 mm. Higher accuracy was generally found in stable midface regions compared to more mobile or lateral facial areas. Deep learning techniques consistently showed superior performance compared to traditional and biomechanical methods. Cone-beam computed tomography (CBCT) was the most frequently used imaging technique, and voxel-based image registration provided the highest accuracy in aligning images. AI models have shown strong potential to significantly improve individualized predictions of soft tissue outcomes in orthognathic surgery. However, due to substantial variability in study methodologies, there is an urgent need for standardized protocols and further multicenter studies to ensure reliability and widespread clinical applicability.
Keywords: Artificial Intelligence, Orthognathic Surgery, Soft Tissue, Deep Learning, CBCT

CITATION INFORMATION: Chookiartsiri, C., Sessirisombat, C., & Sirisoontorn, I. (2025). Effectiveness of Artificial Intelligence in Predicting Soft Tissue Outcomes in Patients Undergoing Orthognathic Surgery: A Systematic Review. *Procedia of Multidisciplinary Research, 3*(7), 47.

INTRODUCTION

Orthognathic surgery is a critical component in the correction of skeletal and dental discrepancies that affect facial harmony, occlusal function, and quality of life. Accurate prediction of both skeletal movements and resulting soft tissue changes is essential for successful treatment planning and patient satisfaction. However, advancements in imaging and digital technologies, conventional planning techniques remain limited by their preciseness on two-dimensional cephalometric analysis, empirical estimation methods, and the clinician's interpretation (Bailey et al., 2004; Proffit, 2013).

Predicting soft tissue changing after surgery remains particularly challenging due to the non-linear and individualized nature of soft tissue profile following skeletal repositioning. The study has shown that even small skeletal movements can produce variable soft tissue changes, making accurate prediction difficult using traditional ratio-based approaches (Gateno et al., 2015). This uncertainty can affect surgical planning, esthetic outcomes, and patient expectations.

The emergence of artificial intelligence (AI) has introduced a transformative opportunity in orthognathic treatment planning. AI, particularly machine learning (ML) and deep learning (DL), has demonstrated significant potential in medical image analysis and predictive modeling. These systems can analyze large datasets to learn complex relationships between preoperative skeletal structures and postoperative outcomes. In orthognathic surgery, AI can assist in automated landmark identification, treatment planning, and simulation of soft tissue changes with improved accuracy and consistency compared to conventional methods (Jeong et al., 2022).

AI-based models have shown superior performance in predicting postoperative soft tissue profiles using three-dimensional imaging data and patient-specific characteristics. For example, convolutional neural networks (CNNs) trained on pre- and postoperative CBCT scans have been able to simulate soft tissue changes with high verity, offering surgeons and patients a more realistic expectation of surgical outcomes (Ter Horst et al., 2021). Moreover, AI enhances interdisciplinary collaboration by providing objective, reproducible simulations that combine orthodontic and surgical planning.

Orthognathic surgery aims not only to correct underlying skeletal discrepancies but also to achieve harmonious facial aesthetics. Changing in facial bones directly affects the overlying soft tissues (such as lips, cheeks, and chin contours), accurate soft tissue prediction has become a crucial component of surgical planning. Accurate soft tissue prediction is essential for optimal treatment planning, enhancing patient communication, and achieving favourable esthetic outcomes (Becker et al., 2013).

However, in recent research, AI has not yet been widely implemented in routine orthognathic workflows. Barriers include the need for strong clinical validation, model transparency, and data standardization. There remains a knowledge gap regarding the clinical effectiveness of AI systems in comparison to conventional techniques, particularly in real-world applications.

The aim of this study is to evaluate how effectively AI can support orthognathic planning and predict soft tissue outcomes. Systematically comparing AI-assisted prediction models with traditional planning approaches, this research aims to provide evidence that may facilitate the integration of AI into surgical workflows, reduce planning errors, and improve esthetic and functional predictability in orthognathic care.

LITERATURE REVIEWS

Artificial Intelligence (AI)

Artificial Intelligence (AI) simulates human intelligence processes, including learning, reasoning, and problem-solving, through computer systems. AI is categorized into narrow AI, specialized for specific tasks such as facial recognition and translation, and general AI, broadly replicating human cognitive functions, although largely theoretical (Peter & Intelligence, 2021).

AI originated in the 1950s with pioneering work by Alan Turing and John McCarthy. Significant advancements have emerged with machine learning (ML), where systems learn from data without explicit programming, and deep learning (DL), employing neural networks to analyze complex patterns in extensive datasets (LeCun et al., 2015).

Core AI components include machine learning algorithms for data-driven learning, deep learning's multi-layered neural networks for image recognition and natural language processing, natural language processing (NLP) capabilities for understanding and generating human language, and computer vision technologies interpreting visual information such as facial recognition.

AI has significantly impacted various sectors. In healthcare, it improves diagnostics and personalized treatment through enhanced medical imaging (Esteva et al., 2017). Education benefits from adaptive learning platforms providing tailored feedback (Holmes et al., 2019). Businesses utilize AI in customer service automation, fraud detection, and recommendation systems. Transportation advancements include autonomous vehicles and traffic management systems (Bojarski et al., 2016), while agriculture employs AI to optimize resource use and predict crop yields.

While AI notably enhances efficiency and personalization, its adoption requires careful management of ethical considerations like bias and transparency. Continuous interdisciplinary research remains essential to harness AI's full potential responsibly.

Orthognathic Surgery

Orthognathic surgery is a corrective surgical procedure addressing skeletal dentofacial deformities, including malocclusion, facial asymmetry, and jaw discrepancies, often performed alongside orthodontics to enhance facial aesthetics and occlusal function (Proffit, 2013). Derived from Greek words meaning "straight jaw," the procedure corrects conditions such as Class II and Class III malocclusions, vertical maxillary excess, mandibular prognathism or retrognathism, and facial asymmetry, involving repositioning of the maxilla, mandible, or both (Bailey et al., 2004). Treatment phases typically include pre-surgical orthodontics for dental alignment, surgical repositioning of skeletal structures, post-surgical orthodontic refinements, and a retention phase to maintain outcomes, usually extending from 18 to 24 months or more (Strunga et al., 2023). Indications for orthognathic surgery include moderate to severe skeletal malocclusions unmanageable by orthodontics alone, significant facial disharmony or asymmetry, functional impairments like obstructive sleep apnea (OSA), masticatory inefficiency, temporomandibular joint (TMJ) problems, and psychological distress related to facial aesthetics (Espeland et al., 2008; Murphy et al., 2011). Surgical techniques commonly employed include Le Fort I osteotomy for maxillary repositioning, bilateral sagittal split osteotomy (BSSO) for mandibular adjustments, genioplasty for chin reshaping, and segmental osteotomies for precise jaw segment repositioning. Recent technological advances, such as 3D imaging, virtual surgical planning (VSP), and computer-aided design/computer-aided manufacturing (CAD/CAM), have enhanced accuracy and predictability (Stokbro et al., 2014). However, challenges such as relapse, accurate soft tissue prediction, potential complications, and psychological adaptation must be carefully managed.

Artificial Intelligence in Orthognathic Surgery: Predicting Clinical Outcomes

Orthognathic surgery is a corrective surgical procedure addressing skeletal dentofacial deformities, including malocclusion, facial asymmetry, and jaw discrepancies, often performed alongside orthodontics to enhance facial aesthetics and occlusal function (Proffit, 2013). Derived from Greek words meaning "straight jaw," the procedure corrects conditions such as Class II and Class III malocclusions, vertical maxillary excess, mandibular prognathism or retrognathism, and facial asymmetry, involving repositioning of the maxilla, mandible, or both

(Bailey et al., 2004). Treatment phases typically include pre-surgical orthodontics for dental alignment, surgical repositioning of skeletal structures, post-surgical orthodontic refinements, and a retention phase to maintain outcomes, usually extending from 18 to 24 months or more (Strunga et al., 2023). Indications for orthognathic surgery include moderate to severe skeletal malocclusions unmanageable by orthodontics alone, significant facial disharmony or asymmetry, functional impairments like obstructive sleep apnea (OSA), masticatory inefficiency, temporomandibular joint (TMJ) problems, and psychological distress related to facial aesthetics (Espeland et al., 2008; Murphy et al., 2011). Surgical techniques commonly employed include Le Fort I osteotomy for maxillary repositioning, bilateral sagittal split osteotomy (BSSO) for mandibular adjustments, genioplasty for chin reshaping, and segmental osteotomies for precise jaw segment repositioning. Recent technological advances, such as 3D imaging, virtual surgical planning (VSP), and computer-aided design/computer-aided manufacturing (CAD/CAM), have enhanced accuracy and predictability (Stokbro et al., 2014). However, challenges such as relapse, accurate soft tissue prediction, potential complications, and psychological adaptation must be carefully managed.

The application of artificial intelligence (AI) in orthognathic surgery is emerging as a transformative approach to enhance the accuracy and predictability of clinical outcomes. AI technologies, particularly machine learning (ML), deep learning (DL), and computer vision, enable detailed analysis of large datasets to accurately predict skeletal movements, occlusal outcomes, facial aesthetics, and postoperative soft tissue changes (Jeong et al., 2022; Schwendicke et al., 2020). Deep learning algorithms using cone-beam computed tomography (CBCT) data can simulate postoperative skeletal and soft tissue positions with superior accuracy compared to traditional methods (Park et al., 2022). AI tools, including convolutional neural networks (CNNs), support the precise prediction of facial profile changes, significantly improving surgical planning and patient counseling (Jeong et al., 2022; Park et al., 2022). Additionally, machine learning classifiers such as support vector machines (SVMs) and random forests aid in predicting surgical stability, relapse risks, and complication probabilities, enhancing decision-making and outcome forecasting (Mohaideen et al., 2022). Personalized AI-driven surgical simulation and planning, integrated with virtual surgical planning (VSP) systems, facilitate interactive visualization and refined treatment strategies, significantly improving treatment acceptance and patient satisfaction (Kato et al., 2023). Despite these advancements, further validation and standardization are crucial to fully integrate AI into orthognathic clinical practice.

Conventional Orthognathic Surgery vs. AI-Assisted Treatment Planning and Outcome Prediction

Orthognathic surgery is a corrective surgical procedure addressing skeletal dentofacial deformities, including malocclusion, facial asymmetry, and jaw discrepancies, often performed alongside orthodontics to enhance facial aesthetics and occlusal function (Proffit, 2013). Derived from Greek words meaning "straight jaw," the procedure corrects conditions such as Class II and Class III malocclusions, vertical maxillary excess, mandibular prognathism or retrognathism, and facial asymmetry, involving repositioning of the maxilla, mandible, or both (Bailey et al., 2004). Treatment phases typically include pre-surgical orthodontics for dental alignment, surgical repositioning of skeletal structures, post-surgical orthodontic refinements, and a retention phase to maintain outcomes, usually extending from 18 to 24 months or more (Strunga et al., 2023). Indications for orthognathic surgery include moderate to severe skeletal malocclusions unmanageable by orthodontics alone, significant facial disharmony or asymmetry, functional impairments like obstructive sleep apnea (OSA), masticatory inefficiency, temporomandibular joint (TMJ) problems, and psychological distress related to facial aesthetics (Espeland et al., 2008; Murphy et al., 2011). Surgical techniques commonly employed include Le Fort I osteotomy for maxillary repositioning, bilateral sagittal split osteotomy (BSSO) for mandibular adjustments, genioplasty for chin reshaping, and segmental osteotomies for precise jaw segment repositioning. Recent technological advances, such as 3D imaging, virtual surgical planning (VSP), and computer-aided design/computer-aided manufacturing (CAD/CAM), have enhanced accuracy and predictability (Stokbro et al., 2014). However, challenges such as relapse, accurate soft tissue prediction, potential complications, and psychological adaptation must be carefully managed.

Orthognathic surgery planning traditionally relies on manual cephalometric analysis, two-dimensional (2D) prediction techniques, plaster models, and clinician experience. While successful, these methods face limitations such as subjectivity, time intensity, and reduced predictive accuracy, especially regarding soft tissue outcomes (Bailey et al., 2004; Proffit, 2013). In contrast, artificial intelligence (AI) has emerged as a powerful alternative, leveraging machine learning (ML), deep learning (DL), and computer vision to enhance precision and efficiency in surgical planning. AI applications include automated landmark identification, 3D surgical simulations, prediction of soft tissue responses, and outcome forecasting (Jeong et al., 2022; Park et al., 2022). Deep learning algorithms, particularly convolutional neural networks (CNNs), have demonstrated superior accuracy over traditional techniques in predicting postoperative soft tissue changes, significantly improving treatment planning and patient counseling (Jeong et al., 2022). AI-powered virtual surgical planning (VSP) platforms facilitate rapid visualization of multiple surgical scenarios, enhancing interdisciplinary communication and patient comprehension (Park et al., 2022; Ter Horst et al., 2021). AI-driven approaches promise improved accuracy, efficiency, and personalization, particularly beneficial in complex or asymmetrical cases. However, ethical considerations, data quality assurance, and clinician training remain crucial for effective integration into clinical practice.

From the literature review, the conceptual framework can be drawn as shown in Figure 1.

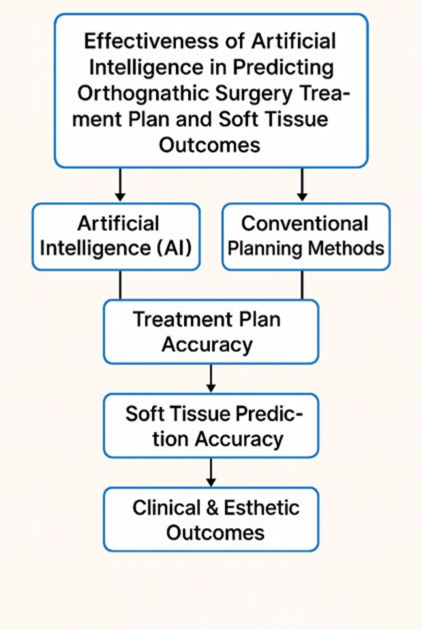

Figure 1 Conceptual Framework

RESEARCH METHODOLOGY

This systematic review was conducted following the Cochrane Handbook guidelines (Chandler et al., 2019) and reported in accordance with the PRISMA statement (Liberati et al., 2009). Electronic databases including PubMed/MEDLINE, Scopus, and ScienceDirect were systematically searched using a predefined strategy from 2014 through April 2025, restricted to English-language articles. Keywords employed included (“Orthognathic surgery”) and (“Artificial intelligence”) and (“Soft tissue”). Eligibility criteria were structured based on

PICOS: Participants comprised patients of all ages, genders, and ethnicities undergoing orthognathic surgery for dentofacial deformities (skeletal Class I, II, or III); interventions involved artificial intelligence (AI) methods such as machine learning, deep learning, or convolutional neural networks for predicting soft tissue outcomes; comparisons were between AI predictions and conventional non-AI methods; primary outcomes included accuracy of AI predictions compared with conventional methods, and secondary outcomes involved postoperative stability, pain, and patient satisfaction. Eligible study designs encompassed randomized controlled trials (RCTs), prospective observational studies, and retrospective comparative studies. Exclusion criteria included studies lacking AI methods, using no CT scans, surgery-first approaches, syndromic or cleft patients, case reports, reviews, conference abstracts, animal or in-vitro studies, or publications not in English. Study selection involved two independent reviewers screening titles and abstracts followed by full-text reviews, with discrepancies resolved by consensus or a third reviewer. Data extraction utilized standardized forms recording study characteristics, participant demographics, intervention details, imaging modalities, comparative methods, primary outcome accuracy assessments, and secondary outcomes such as clinical applicability and patient satisfaction. A descriptive synthesis systematically summarized AI methodologies, accuracy of soft tissue predictions, comparisons with conventional methods, study quality, and clinical implications. Methodological quality was assessed using the Cochrane Risk of Bias tool (RoB 2) for RCTs and the Risk of Bias in Non-Randomized Studies of Interventions (ROBINS-I) for observational studies, with findings reported narratively and in tabular form, highlighting implications for clinical decision-making and the interpretation of results.

RESEARCH RESULTS

Data Collection Process

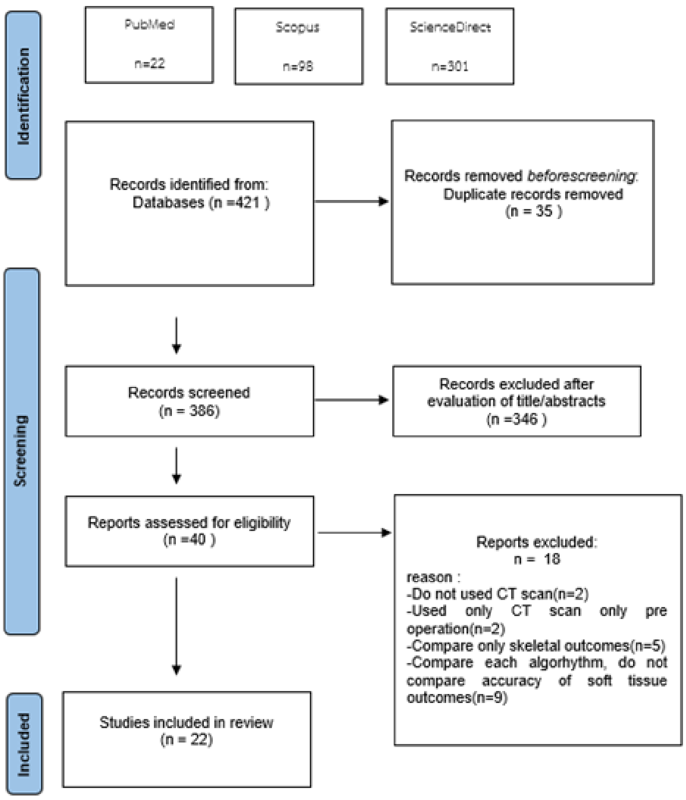

Figure 2 PRISMA Flow Diagram

This systematic review adhered to PRISMA guidelines and involved a comprehensive literature search across Scopus, PubMed, and ScienceDirect, identifying 421 records (98 from Scopus, 22 from PubMed, and 301 from ScienceDirect). After removing 35 duplicates, 386 unique records were screened by two independent reviewers. Title and abstract screening excluded 346 irrelevant articles, leaving 40 for full-text assessment based on PICOS criteria. Eighteen studies were excluded due to reasons such as lack of CT use (n = 2), preoperative CT only (n = 2), skeletal outcome focus (n=5), and absence of soft tissue prediction accuracy evaluation (n = 9). Ultimately, 22 studies were included in the qualitative synthesis, focusing on the effectiveness of AI-based models in predicting soft tissue outcomes in orthognathic surgery (Figure 2).

Data Synthesis

From Table 1, following full-text screening, 22 studies met the inclusion criteria and were included in this systematic review. These studies evaluated the accuracy of three-dimensional (3D) soft tissue simulations by comparing predicted outcomes—generated through virtual treatment planning (VTP)—with actual postoperative results. A variety of commercial and engineering-level software platforms were employed for simulations. The included studies featured sample sizes ranging from 6 to 133 patients, with differing numbers of simulations based on study design.

Most studies focused on skeletal Class III malocclusion (Awad et al., 2022; Hou et al., 2023; Shi et al., 2022), while a smaller number investigated Class II cases or unspecified classifications (Tseng et al., 2024; Yamashita et al., 2022). Simulated procedures commonly included bimaxillary surgery with or without genioplasty, Le Fort I osteotomy (Willinger et al., 2021; Shi et al., 2022), and mandibular osteotomy (Liebregts et al., 2015; Tseng et al., 2024). A few studies included additional soft tissue procedures, though most did not simulate their effects.

CBCT in the upright position was the most widely used imaging modality, followed by MSCT and hybrid imaging approaches integrating CBCT with 3D photography or MRI. Postoperative follow-up imaging was typically acquired between 4 and 12 months, though some studies did not report the interval.

Most studies (n = 19) used actual postoperative skeletal movements for simulations, while three relied on the initial virtual plan, potentially compromising accuracy due to surgical deviations. Accuracy metrics included mean absolute error, root mean square error (RMSE), and percentages of landmarks within 1-3 mm error margins. Reported mean errors ranged from 0.55 mm (Ruggiero et al., 2023) to 2.91 mm (Resnick et al., 2017), with highest accuracy observed in midline regions such as the nose and philtrum, and greatest errors in mobile or lateral regions like the chin, lower lip, and gonial angle.

Simulation techniques varied widely. Commercial tools such as Dolphin 3D and ProPlan CMF used landmark-based algorithms, while biomechanical models (e.g., FEM and MTM) were applied in studies like Alcañiz et al. (2021) and Knoops et al. (2019). Deep learning approaches, including ACMT-Net and Gaussian Mixture Models (GMM), were implemented in Fang et al. (2024) and Tanikawa and Yamashiro (2021), showing promising improvements in prediction accuracy.

Overall, the heterogeneity in software, imaging protocols, and validation methods among studies limits direct comparisons. This highlights the need for standardized protocols and multicenter validation to enhance the consistency and clinical utility of AI-driven soft tissue simulation in orthognathic surgery.

Table 1 Descriptive data on the studies in the review

Author, Year	Study Design	Sample Size	Key Outcomes	Software	Type of Surgery
Alcañiz et al.,2021	Retrospective study	10	Soft-tissue simulation model showed good agreement with clinical outcomes; improved planning accuracy.	FEM	LFI, LFII, BSSRO, Bimax
Awad et al., 2022	Prospective study	20	Soft-tissue prediction accuracy varied across facial regions; higher accuracy in upper facial units (nose, upper cheek, upper lip) compared to lower units (lower cheek, lower lip, chin).	IPS CaseDesigner® (KLS Martin Group) with Geomagic Control X™	Bimax
Fang et al., 2024	Retrospective study	6	Proposed ACMT-Net with a novel CPSA module to predict facial appearance changes based on bony movements; achieved comparable accuracy to FEM-based methods with significantly improved computational efficiency.	ACMT-Net incorporating CPSA module, Deep learning	Bimax
Gutiérrez Venturini et al., 2022	Retrospective study	10	Quantitative analysis measured errors at 9 anatomical landmarks, with greater errors at gonions and lower lip. Qualitative evaluation involved surgeon assessments. No correlation between quantitative and qualitative validations.	FEM-based simulation pipeline	LFI, BSSRO, Bimax
Hou et al., 2023	Retrospective study	58	Found that predictions for upper/lower lips, chin, and buccal regions had errors exceeding 2.0 mm. Some predicted positions higher than actual postoperative positions.	ProPlan CMF software	Bimax
Knoops et al., 2019	Retrospective study	7	Compared 3 methods in Le Fort I osteotomy patients. RMS errors: Dolphin 1.8 ± 0.8 mm, ProPlan 1.2 ± 0.4 mm, PFEM 1.3 ± 0.4 mm. PFEM allowed patient-specific material properties.	Dolphin 3D Ver 11.95, ProPlan CMF Ver 3.0.1, PFEM	LFI
Liebregts et al., 2015	Retrospective study	60	Validated predictability of simulation algorithm for BSSO. Least accuracy in lower lip area. No correlation with amount of mandibular advancement or patient demographics.	MTM algorhithm, Maxilim 2.2.2.1 (Medicim NV, Mechelen, Belgium)	Bimax
Mundluru et al., 2017	Retrospective study	13	3D prediction accuracy of soft tissue changes following surgical correction of facial asymmetry.	Maxilim software	LFI, BSSRO
Ohayon et al., 2025	Retrospective study	27	Improved prediction accuracy for soft tissue outcomes in orthognathic surgery	Cherry imaging, DolphinV11.95. 08.67 SP3	LFI, BSSRO, Bimax
Olivetti et al., 2025	Retrospective study	17	Developed method to predict soft tissue displacement; 70% of points within 2.5 mm error. Significant error at gonion and cheilion.	MATLAB	LFI, BSSRO
Resnick et al., 2017	Retrospective pilot study	7	Mean linear prediction error across 14 facial landmarks: 2.91 ± 2.16 mm. Nasolabial angle error: 8.1 ± 5.6°.	Dolphin 3D software	Le Fort I osteotomy

Author, Year	Study Design	Sample Size	Key Outcomes	Software	Type of Surgery
Ruggiero et al., 2023	Prospective pilot study	8	Patient-specific models using CBCT and MRI. Midface simulation error: 0.55 mm.	FEM	BSSRO
Şenyürek et al., 2023	Retrocpective study	16	Prediction error increased with extent of maxillary advancement. Highest error in mouth region (1.49±0.77 mm).	ProPlan CMF (Materialise)	Le Fort I osteotomy
Shi et al., 2022	Retrospective study	30	Postoperative changes in skeletal Class III patients. Upper lip elongated by 0.71 mm. Soft-to-hard ratios: 0.73:1 (A-Sn), 0.86:1 (Pog-Pg).	CMF ProPlan 3.0 (Materialise)	Bimaxillary surgery (Le Fort I and BSSRO)
Shobair et al., 2021	retrospective study	7	Prediction error across 13 landmarks: 2.78 mm. Better accuracy in midline than lateral regions.	Dolphin 3D	LFI, BSSRO
Tanikawa & Yamashiro, 2021	Retrospective study	72	Developed two AI systems. Average error: 0.94 mm (System S), 0.69 mm (System E). 100% success within 2 mm.	GMM and Deep learning	Orthognathic & orthodontic treatment
Ter Horst et al., 2021	Retrospective study	133 (119 training, 14 test)	Model for 3D profile prediction post mandibular advancement. MAE: 1.0 ± 0.6 mm. Outperformed mass tensor model.	Deep learning Model and MTM algorithm	Mandibular advancement surgery
Tseng et al., 2024	Retrospective study	56	Soft tissue thickness change in hyperdivergent vs. normodivergent groups. Correlation with bone movements.	Dolphin® 11.0 software	BSSRO
Varazzani et al.	Retrospective study	25	Planned vs. postoperative maxillary positions. Lin's concordance >0.95. Discrepancy >1.5 mm in 84% (sagittal/pitch).	Custom 3D VSP software	Bimaxillary non-segmented orthognathic surgery
Verhelst et al., 2021	Retrospective study	15 CBCT scans	Accurate mandibular segmentation. IoU: 94.6%. Reduced segmentation time from 1218s to 17s.	Layer deep learning AI model	Not applicable (segmentation only)
Willinger et al., 2021	Retrospective study	19	Compared Dolphin and IPS for Le Fort II. IPS more accurate at infraorbital rim and sinus floor.	Dolphin Imaging 11.95 and IPS Case Design	Le Fort II osteotomy
Yamashita et al., 2022	Retrospective study	88	3D prediction accuracy acceptable. Mean difference <2 mm in most facial landmarks.	Dolphin 3D software	LFI, BSSRO

DISCUSSION & CONCLUSION

This systematic review offers an in-depth examination of the current capabilities, limitations, and methodological variations in the application of artificial intelligence (AI) for three-dimensional (3D) soft tissue prediction in orthognathic surgery. Across the 22 included studies, the integration of AI into surgical simulation showed promising potential, especially for enhancing the accuracy and personalization of postoperative facial outcome forecasting. However, notable heterogeneity in study designs, image acquisition protocols, simulation algorithms, and validation techniques limits the generalizability and clinical adoption of these tools.

One of the core findings of this review is that while most AI models demonstrated clinically acceptable predictive accuracy—generally within 2 mm—in central midfacial regions such as the nose and philtrum, their performance significantly declined in lateral and mobile facial areas including the lower lip, chin, gonial angles, and cheeks (Hou et al., 2023; Yamashita et al., 2022). These areas are known for their high inter-individual variability and nonlinear soft tissue behavior, which are difficult to capture using simplified prediction models. This regional disparity in accuracy suggests that while current models are sufficient for estimating outcomes in some facial zones, they remain inadequate for more dynamic regions that contribute significantly to facial expression and patient esthetics.

Furthermore, the review identified considerable inconsistencies in image acquisition protocols that likely influenced simulation accuracy. Studies utilizing cone-beam computed tomography (CBCT) in an upright seated or standing position, with relaxed lip posture and centric relation occlusion, achieved better soft tissue simulation fidelity (Awad et al., 2022; Ter Horst et al., 2021). In contrast, studies that employed supine positioning or failed to report imaging posture introduced greater variability, with gravitational soft tissue distortion being a notable concern (Olejnik et al., 2024). The timing of postoperative imaging also played a critical role; early scans taken within 4 to 6 months after surgery may include residual swelling, thereby confounding soft tissue validation (van der Vlis et al., 2014). Only a minority of studies delayed imaging until after 12 months, which is considered optimal for soft tissue stabilization.

Another key issue discussed in this review is the reliability of virtual treatment planning (VTP) as a reference for evaluating simulation accuracy. Several studies, including those by Resnick et al. (2017) and Shobair et al. (2021), assessed prediction performance based solely on the preoperative VTP without accounting for intraoperative deviations or postoperative relapse. This practice introduces bias and underestimates actual simulation errors. Studies by Knoops et al. (2019) and Hou et al. (2023) demonstrated that using postoperative skeletal data rather than planned movements provides a more accurate and clinically relevant benchmark, especially in cases where surgical execution diverges from the initial plan.

The type of soft tissue simulation algorithm employed across studies also influenced predictive performance. Traditional landmark-based systems—commonly embedded in commercial software like Dolphin 3D and ProPlan CMF—were associated with higher error rates in mobile regions due to their linear assumptions and limited capacity to model complex tissue behavior (Resnick et al., 2017; Shobair et al., 2021). Biomechanical models, such as finite element modeling (FEM) and its variants, provided better anatomical fidelity by accounting for material properties and load distribution, but they required substantial preprocessing and assumptions about soft tissue elasticity (Liebregts et al., 2015; Knoops et al., 2019). Deep learning models, particularly convolutional neural networks (CNNs), emerged as the most promising due to their ability to learn complex nonlinear relationships from large datasets. For example, Fang et al. (2024) developed ACMT-Net, which outperformed both FEM and landmark-based systems in both accuracy and computational efficiency. However, the lack of publicly available, annotated datasets and limited model transparency (explainability) remain barriers to clinical adoption (Tanikawa & Yamashiro, 2021).

Rigid registration techniques also varied significantly across studies, affecting alignment precision between simulated and actual outcomes. Voxel-based registration (VB), which aligns 3D datasets using grayscale intensity patterns, was found to offer the highest reproducibility and minimal operator bias (Mundluru et al., 2017; Liebregts et al., 2015). Landmark- and surface-based methods, although more commonly used due to ease of implementation, were more susceptible to user-dependent error, especially when postoperative landmarks were not clearly defined or when craniofacial anatomy had been significantly altered.

In addition to simulation and registration methods, many studies overlooked the impact of adjunctive soft tissue procedures—such as alar cinch sutures, V-Y closures, and mentalis muscle suspension—which are routinely performed to enhance esthetic outcomes. Their omission from simulation models likely contributed to the underestimation of postoperative deviations, particularly around the nose, lips, and chin. Studies that did report the use of such procedures (Mundluru et al., 2017; Resnick et al., 2017; Yamashita et al., 2022) rarely included them in the simulation process, underscoring a major limitation in model realism.

Lastly, postprocessing and evaluation strategies were highly inconsistent across the reviewed studies. Although most studies reported mean and RMS errors, few performed region-specific analyses or employed advanced statistical descriptors to capture variation in soft tissue response. Visual tools like color-coded deviation maps were also underutilized, despite their utility for clinical interpretation and patient communication (Shi et al., 2022). The lack of standardized metrics and inconsistent reporting further complicate cross-study comparisons and reduce the reproducibility of results.

In summary, while AI-driven 3D soft tissue simulations in orthognathic surgery have demonstrated encouraging accuracy in specific facial regions, significant variability in methodology, validation, and clinical integration persists. To enhance reliability and applicability, future research should adopt standardized imaging protocols, incorporate postoperative skeletal outcomes, integrate adjunctive procedures into models, and apply advanced AI algorithms trained on multicenter datasets. Only through these improvements can AI-based simulations achieve routine clinical utility and improve patient-specific surgical outcomes.

REFERENCES

Alcañiz, P., Pérez, J., Gutiérrez, A., Barreiro, H., Villalobos, Á., Miraut, D., ... & Otaduy, M. A. (2021). Soft-tissue simulation for computational planning of orthognathic surgery. *Journal of Personalized Medicine, 11*(10), 982.

Awad, D., Reinert, S., & Kluba, S. (2022). Accuracy of three-dimensional soft-tissue prediction considering the facial aesthetic units using a virtual planning system in orthognathic surgery. *Journal of Personalized Medicine, 12*(9), 1379.

Bailey, L., Cevidanes, L. H., & Proffit, W. R. (2004). Stability and predictability of orthognathic surgery. *Am J Orthod Dentofacial Orthop, 126*(3), 273-277.

Becker, O. E., Scolari, N., Melo, M., Haas Junior, O., Avelar, R., De Menezes, L., & De Oliveira, R. (2013). Three-dimensional planning in orthognathic surgery using cone-beam computed tomography and computer software. *J Comput Sci Syst Biol, 6*(6), 311-316.

Bojarski, M., Del Testa, D., Dworakowski, D., Firner, B., Flepp, B., Goyal, P., Jackel, L. D., Monfort, M., Muller, U., & Zhang, J. (2016). End to end learning for self-driving cars. *arXiv preprint arXiv:1604.07316*.

Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. (2019). *Cochrane handbook for systematic reviews of interventions*. Hoboken: Wiley.

Cho, S. J., Moon, J. H., Ko, D. Y., Lee, J. M., Park, J. A., Donatelli, R. E., & Lee, S. J. (2024). Orthodontic treatment outcome predictive performance differences between artificial intelligence and conventional methods. *The Angle Orthodontist, 94*(5), 557-565.

Eidson, L., Cevidanes, L. H., de Paula, L. K., Hershey, H. G., Welch, G., & Rossouw, P. E. (2012). Three-dimensional evaluation of changes in lip position from before to after orthodontic appliance removal. *American journal of orthodontics and dentofacial orthopedics, 142*(3), 410-418.

Espeland, L., Hogevold, H. E., & Stenvik, A. (2008). A 3-year patient-centred follow-up of 516 consecutively treated orthognathic surgery patients. *Eur J Orthod, 30*(1), 24-30.

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature, 542*(7639), 115-118.

Fang, X., Kim, D., Xu, X., Kuang, T., Lampen, N., Lee, J., Deng, H. H., Liebschner, M. A. K., Xia, J. J., Gateno, J., & Yan, P. (2024). Correspondence attention for facial appearance simulation. *Medical Image Analysis, 93*, Article 103094.

Fang, X., Xiong, X., Lin, J., Wu, Y., Xiang, J., & Wang, J. (2023). Machine-learning-based detection of degenerative temporomandibular joint diseases using lateral cephalograms. *American journal of orthodontics and dentofacial orthopedics, 163*(2), 260-271.e265.

Gateno, J., Alfi, D., Xia, J. J., & Teichgraeber, J. F. (2015). A Geometric Classification of Jaw Deformities. *J Oral Maxillofac Surg, 73*(12 Suppl), S26-31.

Gutiérrez Venturini, A., Guiñales Díaz de Cevallos, J., del Castillo Pardo de Vera, J. L., Alcañiz Aladrén, P., Illana Alejandro, C., & Cebrián Carretero, J. L. (2022). A Quantitative and Qualitative Clinical Validation of Soft Tissue Simulation for Orthognathic Surgery Planning. *Journal of Personalized Medicine, 12*(9), 1460.

Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education promises and implications for teaching and learning*. Center for Curriculum Redesign.

Hou, L., He, Y., Yi, B., Wang, X., Liu, X., Zhang, Y., & Li, Z. (2023). Evaluation of soft tissue prediction accuracy for orthognathic surgery with skeletal class III malocclusion using maxillofacial regional aesthetic units. *Clinical Oral Investigations, 27*(1), 173-182.

Jeong, S. H., Woo, M. W., Shin, D. S., Yeom, H. G., Lim, H. J., Kim, B. C., & Yun, J. P. (2022). Three-Dimensional Postoperative Results Prediction for Orthognathic Surgery through Deep Learning-Based Alignment Network. *J Pers Med, 12*(6).

Kato, R. M., Parizotto, J. D. O. L., Oliveira, P. H. J. D., & Gonçalves, J. R. (2023). Artificial intelligence in orthognathic surgery-a narrative review of surgical digital tools and 3D orthognathic surgical planning. *Journal of the California Dental Association, 51*(1), 2202444.

Khambay, B., & Ullah, R. (2015). Current methods of assessing the accuracy of three-dimensional soft tissue facial predictions: technical and clinical considerations. *International Journal of Oral and Maxillofacial Surgery, 44*(1), 132-138.

Knoops, P., Borghi, A., Breakey, R., Ong, J., Jeelani, N., Bruun, R., Schievano, S., Dunaway, D., & Padwa, B. (2019). Three-dimensional soft tissue prediction in orthognathic surgery: a clinical comparison of Dolphin, ProPlan CMF, and probabilistic finite element modelling. *International Journal of Oral and Maxillofacial Surgery, 48*(4), 511-518.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436-444.

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *Bmj, 339*.

Liebregts, J. H., Timmermans, M., De Koning, M. J., Bergé, S. J., & Maal, T. J. (2015). Three-dimensional facial simulation in bilateral sagittal split osteotomy: a validation study of 100 patients. *Journal of Oral and Maxillofacial Surgery, 73*(5), 961-970.

Mohaideen, K., Negi, A., Verma, D. K., Kumar, N., Sennimalai, K., & Negi, A. (2022). Applications of artificial intelligence and machine learning in orthognathic surgery: A scoping review. *J Stomatol Oral Maxillofac Surg, 123*(6), e962-e972.

Mundluru, T., Almukhtar, A., Ju, X., & Ayoub, A. (2017). The accuracy of three-dimensional prediction of soft tissue changes following the surgical correction of facial asymmetry: An innovative concept. *International Journal of Oral and Maxillofacial Surgery, 46*(11), 1517-1524.

Murphy, C., Kearns, G., Sleeman, D., Cronin, M., & Allen, P. (2011). The clinical relevance of orthognathic surgery on quality of life. *International journal of oral and maxillofacial surgery, 40*(9), 926-930.

Nadjmi, N., Tehranchi, A., Azami, N., Saedi, B., & Mollemans, W. (2013). Comparison of soft-tissue profiles in Le Fort I osteotomy patients with Dolphin and Maxilim softwares. *American Journal of Orthodontics and Dentofacial Orthopedics, 144*(5), 654-662.

Ohayon, C., Bilder, A., Capucha, T., Naki, M., Ginini, J. G., Shilio, D., Gross, N., Rachmiel, A., & Emodi, O. (2025). Accuracy of Three-Dimensional Soft Tissue Prediction in Orthognathic Cases Using Three-Dimensional soft tissue scan implemented in 3D Surgical Planning Software. *Journal of Cranio-Maxillofacial Surgery*.

Olejnik, A., Verstraete, L., Croonenborghs, T. M., Politis, C., & Swennen, G. R. (2024). The Accuracy of Three-Dimensional Soft Tissue Simulation in Orthognathic Surgery—A Systematic Review. *Journal of Imaging, 10*(5), 119.

Olivetti, E. C., Marcolin, F., Moos, S., Vezzetti, E., Borbon, C., Zavattero, E., & Ramieri, G. (2025). How to predict the future face? A 3D methodology to forecast the aspect of patients after orthognathic surgeries. *Computer Methods and Programs in Biomedicine, 265*, 108757.

Park, Y., Choi, J., Kim, Y., Choi, S., Lee, J., Kim, K., & Chung, C. (2022). Deep learning-based prediction of the 3D postorthodontic facial changes. *Journal of Dental Research, 101*(11), 1372-1379.

Peter, N., & Intelligence, R. S. A. (2021). *A Modern Approach*. Pearson Education, USA.

Proffit, W. R. (2013). *Contemporary Treatment of Dentofacial Deformity*. Mosby.

Resnick, C., Dang, R., Glick, S., & Padwa, B. (2017). Accuracy of three-dimensional soft tissue prediction for Le Fort I osteotomy using Dolphin 3D software: a pilot study. *International Journal of Oral and Maxillofacial Surgery, 46*(3), 289-295.

Ruggiero, F., Borghi, A., Bevini, M., Badiali, G., Lunari, O., Dunaway, D., & Marchetti, C. (2023). Soft tissue prediction in orthognathic surgery: Improving accuracy by means of anatomical details. *PLoS ONE, 18*(11 November), Article e0294640.

Schendel, S. A., Jacobson, R., & Khalessi, S. (2013). 3-dimensional facial simulation in orthognathic surgery: is it accurate?. *Journal of Oral and Maxillofacial Surgery, 71*(8), 1406-1414.

Schwendicke, F. a., Samek, W., & Krois, J. (2020). Artificial intelligence in dentistry: chances and challenges. *Journal of Dental Research, 99*(7), 769-774.

Şenyürek, S. A., Ajami, S., Ruggiero, F., Van de Lande, L., Caron, C. J., Schievano, S., Dunaway, D. J., Padwa, B., Koudstaal, M. J., & Borghi, A. (2023). The accuracy of computer-assisted surgical planning in predicting soft tissue responses after Le Fort I osteotomy: retrospective analysis. *Journal of Craniofacial Surgery, 34*(1), 131-138.

Shi, Y., Liu, S., Shao, X., Zong, C., Bai, S., Yang, Y., Liu, Y., Shang, H., & Tian, L. (2022). Facial changes in patients with skeletal class III deformity after bimaxillary surgery: an evaluation based on three-dimensional photographs registered with computed tomography. *British Journal of Oral and Maxillofacial Surgery, 60*(10), 1404-1410.

Shobair, N., Ghanem, A. A., Ayoub, A., Barakat, A., & Diaa, M. (2021). The Prediction Accuracy of soft tissue changes following orthognathic surgery using Dolphin 3D Software package; A pilot study. *Ain Shams Dental Journal (Egypt), 23*(3), 55-61.

Stokbro, K., Aagaard, E., Torkov, P., Bell, R., & Thygesen, T. (2014). Virtual planning in orthognathic surgery. *International journal of oral and maxillofacial surgery, 43*(8), 957-965.

Stratemann, S., Huang, J., Maki, K., Miller, A., & Hatcher, D. (2008). Comparison of cone beam computed tomography imaging with physical measures. *Dentomaxillofacial Radiology, 37*(2), 80-93.

Strunga, M., Urban, R., Surovková, J., & Thurzo, A. (2023). Artificial intelligence systems assisting in the assessment of the course and retention of orthodontic treatment. In *Healthcare* (Vol. 11, No. 5, p. 683). MDPI.

Tanikawa, C., & Yamashiro, T. (2021). Development of novel artificial intelligence systems to predict facial morphology after orthognathic surgery and orthodontic treatment in Japanese patients. *Sci Rep, 11*(1), 15853.

Ter Horst, R., van Weert, H., Loonen, T., Bergé, S., Vinayahalingam, S., Baan, F., Maal, T., de Jong, G., & Xi, T. (2021). Three-dimensional virtual planning in mandibular advancement surgery: Soft tissue prediction based on deep learning. *J Craniomaxillofac Surg, 49*(9), 775-782.

Tseng, Y. C., Wu, T. Y., Lu, C. Y., Chou, S. T., Lin, S. H., & Chen, C. M. (2024). Investigating the postoperative soft tissue changes in different vertical facial divergent patients with mandibular prognathism. *Journal of Dental Sciences, 19*(3), 1443-1451.

Ullah, R., Turner, P., & Khambay, B. (2015). Accuracy of three-dimensional soft tissue predictions in orthognathic surgery after Le Fort I advancement osteotomies. *British Journal of Oral and Maxillofacial Surgery, 53*(2), 153-157.

Van Der Vlis, M., Dentino, K. M., Vervloet, B., & Padwa, B. L. (2014). Postoperative swelling after orthognathic surgery: a prospective volumetric analysis. *Journal of Oral and Maxillofacial Surgery, 72*(11), 2241-2247.

Varazzani, A., Tognin, L., Corre, P., Bouletreau, P., Perrin, J. P., Menapace, G., Bergonzani, M., Pedrazzi, G., Anghinoni, M., & Poli, T. (2025). Virtual surgical planning in orthognathic surgery: A prospective evaluation of postoperative accuracy. *Journal of Stomatology, Oral and Maxillofacial Surgery, 126*(1), 102025.

Verhelst, P.-J., Smolders, A., Beznik, T., Meewis, J., Vandemeulebroucke, A., Shaheen, E., Van Gerven, A., Willems, H., Politis, C., & Jacobs, R. (2021). Layered deep learning for automatic mandibular segmentation in cone-beam computed tomography. *Journal of Dentistry, 114*, 103786.

Willinger, K., Guevara-Rojas, G., Cede, J., Schicho, K., Stamm, T., & Klug, C. (2021). Accuracy of soft tissue prediction of 2 virtual planning systems in patients undergoing intraoral quadrangular Le Fort II osteotomy. *Plastic and Reconstructive Surgery-Global Open, 9*(2), e3326.

Yamashita, A. L., Iwaki Filho, L., Ferraz, F. W. D. S., Ramos, A. L., Previdelli, I. T. D. S., Pereira, O. C. N., ... & Iwaki, L. C. V. (2022). Accuracy of three-dimensional soft tissue profile prediction in orthognathic surgery. *Oral and Maxillofacial Surgery*, 1-9.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Conflicts of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.