

Orthognathic vs. Orthodontic Treatment: A Comparative Study of AI and Clinical Diagnosis

*2024 Orthodontic Faculty Development
Fellowships (OFDFA)*

Dr Jonas Bianchi Bianchi

jbianchi@pacific.edu
O: 734-846-9793

FollowUp Form

Award Information

In an attempt to make things a little easier for the reviewer who will read this report, please consider these two questions before this is sent for review:

- Is this an example of your very best work, in that it provides sufficient explanation and justification, and is something otherwise worthy of publication? (We do publish the Final Report on our website, so this does need to be complete and polished.)*
- Does this Final Report provide the level of detail, etc. that you would expect, if you were the reviewer?*

Title of Project*

Orthognathic vs. Orthodontic Treatment: A Comparative Study of AI and Clinical Diagnosis

Award Type

Orthodontic Faculty Development Fellowship Award (OFDFA)

Period of AAOF Support

July 1, 2024 through June 30, 2025

Institution

University of the Pacific, Arthur A. Dugoni School of Dentistry

Names of principal advisor(s) / mentor(s), co-investigator(s) and consultant(s)

Lucia Cevidanes, Heesoo Oh, Joorok Park, Juan Prieto, Joao Goncalves

Amount of Funding

\$30,000.00

Abstract

(add specific directions for each type here)

I am honored to have been granted the Orthodontic Faculty Development Fellowship (OFDFA) for the second consecutive year, an opportunity that has significantly transformed my career as an educator. The past awards have played a pivotal role in my academic growth, reflected in my increased publication record and my first independent research project application with the NIH R16 award called: "KIDS-SAFE - An Innovative

AI-Driven System for Radiation-Free Craniofacial Asymmetry Characterization and Prediction in Children," which is currently under review. As a highly motivated junior faculty member, the OFDFA 2024 can continue advancing my career development in teaching, education, clinical practice, and research. This progress is made possible through a meticulously crafted mentoring plan, with the invaluable guidance of a dedicated team of mentors and consultants. My objectives in education and teaching include active involvement in departmental, university, and national-level educational experiences. This includes participating in specialized training courses offered by the University of the Pacific and engaging in weekly conferences with Dr. Heesoo Oh, Chair of UoP's Orthodontics department, to enhance my didactic teaching in the graduate course. Additionally, I plan to contribute actively as a member of the Angle Society and continue serving as a peer reviewer for various Orthodontic journals. On the clinical aspect, my development plan involves integrating the advancements derived from my proposed research into practical implementation to enhance decision-making in orthodontic patient treatment. In the research area, this study will evaluate the agreement and predictive performance among two orthodontists and two oral maxillofacial surgeons in determining the need for orthognathic surgery using cephalometric data from lateral radiographs. This will involve assessing both intra- and inter-agreement among clinicians and comparing their assessments with those generated by state-of-the-art machine learning models supported by rigorous hypothesis testing. Overall, the Orthodontic Faculty Development Fellowship Award from AAOF provides critical support that enables me to concentrate on these essential educational, clinical, and research objectives as a dedicated orthodontic faculty member. My long-term vision is to disseminate the patient-specific tools and insights gained from this OFDFA-supported project to benefit clinicians and residents, ultimately contributing to improved patient care and outcomes.

Respond to the following questions:

Detailed results and inferences:*

If the work has been published, please attach a pdf of manuscript below by clicking "Upload a file".

OR

Use the text box below to describe in detail the results of your study. The intent is to share the knowledge you have generated with the AAOF and orthodontic community specifically and other who may benefit from your study. Table, Figures, Statistical Analysis, and interpretation of results should also be attached by clicking "Upload a file".

reporttogether-compressed.pdf

Yes, the work has been published and presented through multiple sources. The findings on clinician agreement were presented at the 2025 AADOCR meeting in New York; both the abstract and the poster from this meeting are attached in this section. In addition, a book chapter was published as part of the Moyers Symposium proceedings. One additional manuscript has also been published as an outcome of this grant, focusing on the application of artificial intelligence in orthodontics and education.

The study revealed that overall agreement among the four clinicians (two orthodontists and two oral maxillofacial surgeons) was low when diagnosing the need for orthognathic surgery versus orthodontic treatment based on cephalometric data. Individual accuracy varied, with Orthodontist A achieving the highest individual accuracy (70%) and Surgeon B the lowest (45%). However, consensus diagnosis among the clinicians significantly improved accuracy, with both Orthodontist A and Surgeon D reaching 86.4% accuracy when decisions were made collaboratively. These results highlight the value of consensus in complex treatment planning and underscore the potential of machine learning models, which demonstrated higher consistency compared to individual clinician decisions.

Were the original, specific aims of the proposal realized?*

Yes, the original specific aims of the proposal were successfully realized.

On the research side, we achieved the primary goals of evaluating intra- and inter-clinician agreement on the need for orthognathic surgery and comparing these assessments with machine learning models. We collected and analyzed data from 920 patients, applied 10 machine learning models, and conducted detailed comparisons of clinician and AI performance. The results were disseminated through presentations (e.g., AADOCR 2025), a book chapter (Moyer Symposium proceedings), and peer-reviewed publications.

On the educational and teaching goals, I met my short-term goals by continuing regular mentorship of residents and dental students, expanding their research engagement, and offering the intensive training on Artificial Intelligence to undergraduate students. I actively participated in faculty development opportunities and integrated research findings into the classroom, enriching student learning with real-world applications of AI in orthodontics.

Regarding clinical teaching goals, I maintained and deepened my collaboration with Dr. Park through co-supervision of clinic days, bi-weekly meetings, and participation in case analysis seminars. I continued mentoring residents on evidence-based patient care and treatment planning, incorporating digital orthodontics and AI tools. Additionally, I pursued opportunities to broaden my clinical exposure by observing and learning from other faculty members' treatment approaches.

Together, these efforts reflect the successful realization of the proposal's original aims in research, education, and clinical teaching, advancing both personal development and contributions to the field of orthodontics.

Were the results published?*

Yes

Have the results of this proposal been presented?*

Yes

To what extent have you used, or how do you intend to use, AAOF funding to further your career?*

The AAOF funding has been instrumental in advancing my faculty, research, and clinical goals. The financial support enabled me to acquire essential equipment, including a high-performance computer that provided the computational capacity necessary to complete my machine learning research on predicting the need for orthognathic surgery. This equipment has also supported other projects beyond the original study, further amplifying the impact of the funding, as reflected in my recent publications.

The grant made it possible for me to sustain my position as a full-time faculty member in San Francisco while pursuing my Advanced Standing Program in Orthodontics in the United States. This support has been especially critical in a high cost-of-living area and has allowed me to maintain strong contributions in research, teaching, and patient care.

The funding also enabled me to engage in specialized training and professional development workshops, significantly strengthening my ability to integrate advanced technologies into education. This has directly

benefited students and residents, as I have incorporated AI tools and digital orthodontics into both coursework and clinical training.

The visibility gained through this work led to invitations to present at several prominent meetings, including the AAO Annual Session, the Idaho Society of Orthodontists, and the Frederick West Lecture in San Francisco. In addition, I applied the research methods developed through this project as part of a successful NIH R01 grant application, where I serve as a subcontract investigator further extending the impact and value of the AAOF's support.

Overall, the AAOF funding has provided essential resources and opportunities that continue to help me grow as an academic orthodontist, improve patient care through innovative research, and contribute to the training of future professionals in the field.

Accounting: Were there any leftover funds?

\$0.00

Published

Citations*

You indicated results have been published. Please list the cited reference/s for publication/s including titles, dates, author or co-authors, journal, issue and page numbers

1. Lowder, B., Toyama, C., Roberto Goncalves, J., Dos Santos Pintoa, A., Gonzaga Gandini Junior, L., Paladini, G., Alves De Oliveria Junior, M., Augusto Campanaha, A., Bianchi, J., Henrique Jose De Oliveira, P., Li, T., Li, H., & Cevidanes, L. (2025). Assessment of clinical decisions between orthodontic and orthognathic surgery treatments [Conference poster]. 2025 AADOCR/CADR Annual Meeting, New York City, NY.
2. Bianchi, J., Li, T., Li, H., Gonçaves, J. R., Gandini Junior, L. G., de Oliveira, P. H. J., & Cevidanes, L. (2025). Leveraging AI and data science for better outcomes in orthodontics and orthognathic surgery. In *The 51st Moyers Symposium Monograph: Artificial intelligence: What should clinicians know?* (Vol. 60). Craniofacial Growth Series, University of Michigan. <https://dx.doi.org/10.7302/25242>
3. Bianchi J, Zheng M. Leveraging Generative Artificial Intelligence in Teaching, Scholarship and Dental Education: Use Cases and Reflections. *Orthod Craniofac Res*. 2025 Jun 6. doi: 10.1111/ocr.12949. Epub ahead of print. PMID: 40476658.

Was AAOF support acknowledged?

If so, please describe:

1. The abstract and poster included direct acknowledgment of the AAOF, stating: "Support Funding Agency/Grant Number: American Association of Orthodontists Foundation Robert E. Binder Teaching Fellowship Award and National Institute of Dental and Craniofacial Research/NIH: R01DE024450."
2. The AAOF was also acknowledged during the lecture presentation associated with the book chapter. Although the published chapter itself did not include a formal acknowledgments section due to its formatting, I made sure to recognize the AAOF during the lecture as the primary supporter and funder of the project.

3. The manuscript includes a direct citation of AAOF support: "Funding: This work was supported by the American Association of Orthodontists Foundation, Robert E. Binder Teaching Fellowship Award."

Presented

Please list titles, author or co-authors of these presentation/s, year and locations:*

1. Meeting: 2025 AADOCR/CADR Annual Meeting (New York City, New York)

Location: New York City, New York

Year: 2025

Abstract Category|Abstract Category(s): Orthodontics Research

Title: Assessment of Clinical Decisions Between Orthodontic and Orthognathic Surgery Treatments.

Authors:

Bianchi, Jonas (University of the Pacific, Arthur A. Dugoni School of Dentistry , San Francisco , California , United States)

Lowder, Brendon (University of the Pacific, Arthur A. Dugoni School of Dentistry , San Francisco , California , United States)

Toyama, Claudia (Private practice , São Paulo , Brazil)

Roberto Goncalves, Joao (São Paulo State University (Unesp), School of Dentistry , São Paulo , Araraquara , Brazil)

Dos Santos Pintoa, Ary (São Paulo State University (Unesp), School of Dentistry , São Paulo , Araraquara , Brazil)

Gonzaga Gandini Junior, Luiz (São Paulo State University (Unesp), School of Dentistry , São Paulo , Araraquara , Brazil)

Paladini, Guilherme (Private practice , São Paulo , Brazil)

Alves De Oliveria Junior, Melchiades (Private practice , São Paulo , Brazil)

Augusto Campanaha, Antonio (Private practice , São Paulo , Brazil)

Henrique Jose De Oliveira, Pedro (São Paulo State University (Unesp), School of Dentistry , São Paulo , Araraquara , Brazil)

Li, Tengfei (University of North Carolina at Chapel Hill , Chapel hill , North Carolina , United States)

Li, Haoyue (University of North Carolina at Chapel Hill , Chapel hill , North Carolina , United States)

Cevidanes, Lucia (University of Michigan , Ann Arbor , Michigan , United States)

2. Meeting: American Association of Orthodontists Annual Session 2025

Location: Philadelphia - PA

Year: 2025

Joint Oral Presentation: AI in Orthodontics: Beyond the Hype - 2025 Annual Session

Authors: Bianchi, Jonas (University of the Pacific, Arthur A. Dugoni School of Dentistry , San Francisco , California , United States)

Cevidanes, Lucia (University of Michigan , Ann Arbor , Michigan , United States)

3. Meeting: 38th Annual Frederick T. West Orthodontic Lectureship

Location: San Francisco - CA

Year: 2025

Oral Presentation: Artificial Intelligence in Orthodontics: Are We Entering a New Normal?

Authors: Bianchi, Jonas (University of the Pacific, Arthur A. Dugoni School of Dentistry , San Francisco , California , United States)

4. Meeting: The 51st Annual Moyers Symposium - The University of Michigan

Location: Ann Arbor, MI.

Year: 2024

Oral Presentation: Empowering Clinicians: Harnessing AI and Data-Driven Tools in Orthognathic Surgery.

Authors: Bianchi, Jonas (University of the Pacific, Arthur A. Dugoni School of Dentistry , San Francisco , California , United States)

Was AAOF support acknowledged?

If so, please describe:

Yes, AAOF support was acknowledged in every oral presentation and lecture I delivered related to this project. During each talk, I made a point to recognize the American Association of Orthodontists Foundation as the primary supporter of my work, emphasizing the vital role their funding played in enabling the research and advancing my academic career.

Internal Review

Reviewer comments

Reviewer Status*

File Attachment Summary

Applicant File Uploads

- reporttogether-compressed.pdf

IADR Abstract Archives

Assessment of Clinical Decisions Between Orthodontic and Orthognathic Surgery Treatments.

Objectives: To assess the performance of four experts in diagnosing orthognathic surgery or conventional orthodontics and compare their individual and collective performance based on the treatment received by the patients and consensus decisions.

Methods: 920 lateral radiographs (558 Class II, 362 Class III) from patients treated with either orthodontics or orthognathic surgery were analyzed. Thirty-two cephalometric measurements were presented to each expert to make a diagnosis. Inclusion criteria were pre-treatment lateral radiographs of patients aged >16, at CS5/CS6 vertebral maturation, with skeletal Class II (ANB >4) or Class III (ANB <0) malocclusions.

Results: Overall, there was low agreement among the four experts. Ortho A had the highest accuracy (0.70) for true treatment, while Surgeon B had the lowest (accuracy = 0.45). Consensus diagnosis improved performance across all experts, with Surgeon D and Ortho A both achieving an accuracy of 0.864.

Conclusions: The consensus diagnosis outperformed individual true treatment predictions, emphasizing the importance of collaborative decision-making in managing Class II and III malocclusions, whether the treatment required was surgical or conventional orthodontics.

Division:

Meeting: 2025 AADOCR/CADR Annual Meeting (New York City, New York)

Location: New York City, New York

Year: 2025

Final Presentation ID: 1471

Abstract Category|Abstract Category(s): Orthodontics Research

Authors

- **Lowder, Brendon** (University of the Pacific, Arthur A. Dugoni School of Dentistry , San Francisco , California , United States)
- Toyama, Claudia (Private practice , São Paulo , Brazil)
- Roberto Goncalves, Joao (São Paulo State University (Unesp), School of Dentistry , São Paulo , Araraquara , Brazil)
- Dos Santos Pintoa, Ary (São Paulo State University (Unesp), School of Dentistry , São Paulo , Araraquara , Brazil)
- Gonzaga Gandini Junior, Luiz (São Paulo State University (Unesp), School of Dentistry , São Paulo , Araraquara , Brazil)
- Paladini, Guilherme (Private practice , São Paulo , Brazil)
- Alves De Oliveria Junior, Melchiades (Private practice , São Paulo , Brazil)
- Augusto Campanaha, Antonio (Private practice , São Paulo , Brazil)
- Bianchi, Jonas (University of the Pacific, Arthur A. Dugoni School of Dentistry , San Francisco , California , United States)
- Henrique Jose De Oliveira, Pedro (São Paulo State University (Unesp), School of Dentistry , São Paulo , Araraquara , Brazil)
- Li, Tengfei (University of North Carolina at Chapel Hill , Chapel hill , North Carolina , United States)
- Li, Haoyue (University of North Carolina at Chapel Hill , Chapel hill , North Carolina , United States)
- Cevidanes, Lucia (University of Michigan , Ann Arbor , Michigan , United States)

Support Funding Agency/Grant Number: American Association of Orthodontists Foundation Robert E. Binder Teaching Fellowship Award and National Institute of Dental and Craniofacial Research/NIH: R01DE024450.

Financial Interest Disclosure: None

SESSION INFORMATION

Poster Session

Advancing Orthodontics: Patient Outcomes, Disparities, and Emerging Research

Saturday, 03/15/2025 , 11:00AM - 12:15PM

BACKGROUND

Orthodontic treatment requires precise measurements of dental and skeletal parameters, including facial aesthetics, occlusion, and functionality. Choosing between orthodontics and orthognathic surgery remains challenging, especially for borderline cases. Technological advancements facilitate clinical decisions, with big data enabling AI-guided diagnosis¹⁻⁴.

The aim of this retrospective study is to use cephalometric data to evaluate ten different machine learning (ML) models and their prediction results for surgery or orthodontics treatment and to assess clinicians' agreements.

Our hypothesis states that the ML models can accurately predict the need for surgery, similar to the clinician's agreement or real treatment is done.



Fig 1. Two patients with skeletal class II. Left: Orthodontically treated only, and right: surgical correction.

MATERIALS AND METHODS

This study was approved by the IRB of the UNESP university. A total of 920 lateral radiographs from patients previously treated with conventional orthodontics or combined with orthognathic surgery were used, comprising n=558 Class II and n=362 Class III patients, respectively. Thirty-two measures were obtained from each cephalogram at the initial appointment. The subjects were randomly divided into training (n=552), validation (n=183), and test (n=185) datasets, both as an entire sample and divided into Class II and Class III subgroups. The extracted data were evaluated using 10 machine learning models and by a four-expert panel consisting of orthodontists (n=2) and surgeons (n=2). Inclusion criteria included pre-treatment lateral radiographs of skeletal Class II (ANB>4 or Witts>0) and Class III (ANB<0 or Witts<0) patients who were candidates for orthognathic surgery or orthodontic treatment, with or without premolar extractions, aged >16 years, and at stages CS5 or CS6 of vertebral maturation.

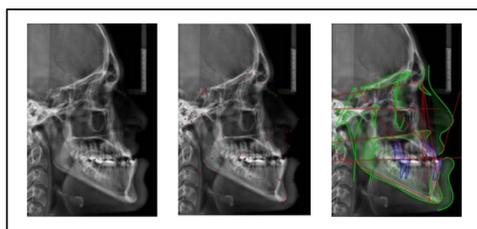


Fig 2. Lateral cephalogram was used to create the cephalometric tracing

Table 1. Cephalometric variables that were measured in lateral radiographs.

Variable	Description
Relation between maxilla and mandible with the cranial base	
SNA	Relation between the maxilla and the cranial base
SNS	Relation between the mandible and the cranial base
A-Nppp	Relation between the maxilla and the cranial base
Pog-Nppp	Relation between the mandible and the cranial base
Cranial base evaluation	
SNSi	Cranial base angle
Ant. Cran. Base	Anterior Cranial Base's length
Post. Cran. Base	Posterior Cranial Base's length
Relation between maxilla and mandible	
ANB	Angular relationship between maxilla and mandible
Wits	Linear relationship between maxilla and mandible
Co-A	Maxilla's length
Co-S	Mandible's length
Vertical Analysis	
S-I (Co-Me)	Cranial base and mandibular plane angle
S-I (Co-N)	the cranial base and Gonion angle
FMA	Frankfort to mandibular plane angle
ML-Me	Frankfort to mandibular plane angle
FM-Me	Frankfort to maxillary plane angle
ML-Me	Frankfort to maxillary plane angle
ML-Me	Lower Anterior Facial Height
ML-Me	Lower posterior Facial Height

Dental Analysis	
1-NA	Upper incisor's inclination
1-NA	Upper incisor's protrusion
1-NB	Lower incisor's inclination
1-NB	Lower incisor's protrusion
IMPA	Mandibular incisor plane angle
FMA	Frankfurt mandibular incisor plane angle
FMA	Mandible's incisor to maxillary plane angle
1-Planoxk	
Profile Analysis	
Pm (Sn-Ls)	Nasolabial angle
H-Nariz	Holdway line distance to pro-nasal point
HNB	Profile's evaluation
Occlusal Plane Analysis	
S-N Plano Oclusal	Cranial base and occlusal plane angle
F Plano Oclusal	Frankfurt plane and occlusal plane angle
Airways Analysis	
Nla-Nlp	Upper airways analysis
Blc-Blp	Lower airways analysis

Fig 3. Cephalometric variables used in the study.

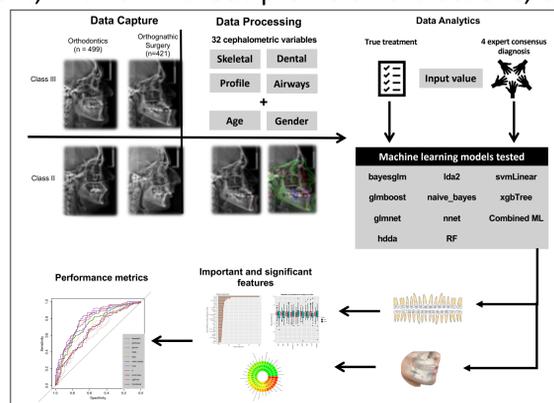


Fig 4. General overview of the study methodology and ML model tests.

RESULTS AND DISCUSSION

True Treatment - Entire Sample							
RF alg	Accuracy	Precision_surgery	Precision_ortho	Recall_surgery	Recall_ortho	F1-score	AUC
Validation	0.714	0.737	0.687	0.737	0.675	0.709	0.781
Testing	0.668	0.676	0.658	0.737	0.588	0.663	0.725

Consensus - Entire Sample							
Combined ML	Accuracy	Precision_surgery	Precision_ortho	Recall_surgery	Recall_ortho	F1-score	AUC
Validation	0.797	0.818	0.784	0.684	0.883	0.788	0.871
Testing	0.707	0.714	0.7	0.667	0.745	0.706	0.791

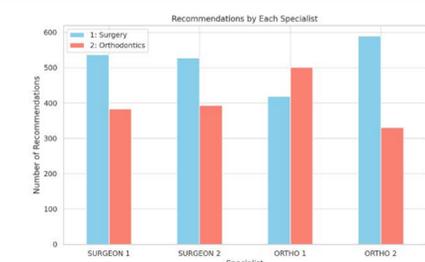


Fig 7. Clinicians' Treatment of Choice per Case

True Treatment - CI II patients							
bayesglm alg	Accuracy	Precision_surgery	Precision_ortho	Recall_surgery	Recall_ortho	F1-score	AUC
Validation	0.664	0.688	0.606	0.803	0.455	0.63	0.703
Testing	0.777	0.769	0.794	0.896	0.6	0.756	0.847

Consensus - CI II patients							
Combined ML	Accuracy	Precision_surgery	Precision_ortho	Recall_surgery	Recall_ortho	F1-score	AUC
Validation	0.855	0.852	0.857	0.881	0.824	0.853	0.905
Testing	0.759	0.722	0.793	0.765	0.754	0.758	0.824

True Treatment - CI III patients							
bayesglm alg	Accuracy	Precision_surgery	Precision_ortho	Recall_surgery	Recall_ortho	F1-score	AUC
Validation	0.789	0.718	0.875	0.875	0.718	0.789	0.78
Testing	0.863	0.897	0.841	0.788	0.925	0.86	0.904

Consensus - CI III patients							
Combined ML	Accuracy	Precision_surgery	Precision_ortho	Recall_surgery	Recall_ortho	F1-score	AUC
Validation	0.873	0.806	0.925	0.893	0.86	0.87	0.923
Testing	0.822	0.909	0.784	0.645	0.952	0.807	0.89

Fig 5. Results of the ML models for the True treatment.

Fig 6. Results of the ML models for the Clinicians Consensus

Boxplot for the most important features for consensus diagnosis in class II (A) and class III patients (B)

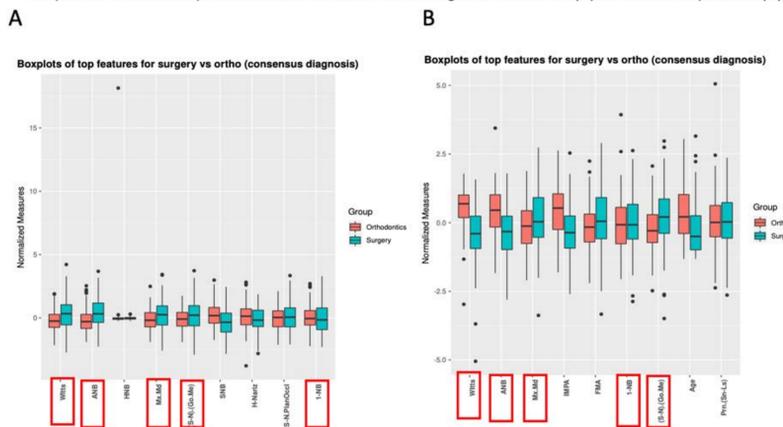


Fig 8 – Boxplot showing the top 10 variables with the highest participation in the decision-making.

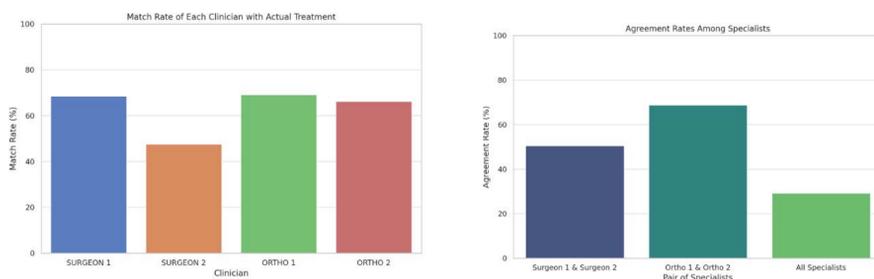


Fig 9. Clinicians' agreement among themselves (right) and with the actual treatment (left)

• **Accuracy** is about the overall correctness of the model.
 • **Precision** is about how accurate the positive predictions are.
 • **Recall** is about how well the model can identify positive cases. "sensitivity"
 • **F1 Score** is a balanced measure that considers both precision and recall.

CONCLUSION

The machine learning methods proposed in this study can accurately predict the need for orthognathic surgery in class II and III patients. The performance was better when consensus diagnosis was used than true treatment because of the lack of subjectiveness and human bias. Witts and ANB were the most important features to predict which would be the best treatment for the patient. Class III patients showed the best results, mainly with the combinate machine learning method for the validation dataset. Also, clinicians have a lower agreement rate, with ML models showing higher consistency in the prediction.

REFERENCES & ACKNOWLEDGEMENTS

- Moyers R. *Ortodontia*. 3. ed. Guanabara Koogan; 1979.
- Reis SAB, Abrão J, Capelozza Filho L, Claro CA de A. Análise Facial Subjetiva. *Rev Dent Press Ortod e Ortop Facial*. 2006;11(5):159-172. doi:10.1590/S1415-54192006000500017
- Arnett GW, Bergman RT. Facial keys to orthodontic diagnosis and treatment planning. Part I. *Am J Orthod Dentofac Orthop*. 1993;103(4):299-312. doi:10.1016/0889-5406(93)70010-L
- William Arnett G, Bergman RT. Facial keys to orthodontic diagnosis and treatment planning—part II. *Am J Orthod Dentofac Orthop*. 1993;103(5):395-411. doi:10.1016/S0889-5406(05)81791-3

LEVERAGING AI AND DATA SCIENCE FOR BETTER OUTCOMES IN ORTHODONTICS AND ORTHOGNATHIC SURGERY

*Jonas Bianchi, Tengfei Li, Haoyue Li, João Roberto Gonçalves,
Luiz Gonzaga Gandini Junior, Pedro Henrique José de Oliveira, Lucia Cevidanes*

ABSTRACT

When skeletal, occlusal, and aesthetic discrepancies are too pronounced to be addressed with compensatory orthodontics, orthognathic surgery becomes necessary to restore the patient's facial harmony and function. However, borderline cases often present significant challenges. The clinician's experience and empirical knowledge are crucial for effective treatment planning in these cases. The growing volume of data and advancements in artificial intelligence (AI) offer new opportunities to enhance decision-making processes. AI models, particularly those utilizing machine learning (ML) and deep learning (DL), can assist orthodontists and maxillofacial surgeons in improving diagnostic accuracy and treatment planning. Practitioners can refine their strategies and outcomes by comparing their clinical assessments with expert opinions derived from extensive datasets. AI is not intended to replace human expertise but to augment it, offering an additional tool to support high-quality diagnostics and treatment plans. This chapter explores the transformative applications of AI in healthcare, specifically in dentistry and orthognathic surgery, highlighting the benefits and challenges of integrating AI technologies in clinical practice.

KEY WORDS: Artificial intelligence, Orthodontics, Orthognathic surgery, Data science

INTRODUCTION

Artificial intelligence (AI) tools have been used in many different areas of healthcare. They are not limited to a particular area, and their usage has grown since the amount of data has increased due to digitalization, more access to different exams, and technological improvements. They have been used by other investigative teams mainly to improve epidemiological studies, trace diseases, give a more objective understanding of exams, and recognize patterns that a human mind may find more challenging to see [1-4]. The use of AI in dentistry does not differ much from its primary use in healthcare. Machine learning (ML) and deep learning (DL) are the most usual subsets of AI used in dentistry. ML and DL are transformative technologies in the field of dentistry, each with distinct applications and advantages [5-8], and are widely used in dentistry for predictive analytics, diagnosis, and personalized treatment plans. ML algorithms analyze historical dental data to predict patient outcomes, such as the likelihood of developing periodontal disease or the success of orthodontic treatments. They help identify patterns in patient data for early diagnosis and risk assessment, recommending preventive measures tailored to individual needs. Additionally, ML optimizes dental practice operations, improving resource allocation, appointment scheduling, and inventory management. One of the key advantages of ML is its interpretability, particularly with models such as decision trees and linear regression, which dental professionals can easily understand and trust. Furthermore, ML can effectively function with smaller datasets, making it useful in areas with limited data. However, ML requires manual effort for feature engineering, which involves selecting and extracting relevant features from raw data, and its performance may be less accurate than DL in complex tasks such as image recognition [6, 9].

DL, a subset of ML, excels in analyzing complex data types such as dental images and 3D scans. DL models, particularly convolutional neural networks (CNNs), are highly effective in dental imaging, detecting caries, identifying fractures, and diagnosing orthodontic anomalies. In addition, DL facilitates the analysis of dental X-rays, cone beam computed tomography (CBCT) scans, and photographs, improving diagnostic accuracy and treatment planning. Natural language processing (NLP) applications of DL extract useful information from unstructured dental records, clinical notes, and research papers. DL also aids in designing custom dental prosthetics by predicting the optimal shape and fit based on patient-specific data. One of the main advantages of DL is its ability to automatically extract features from raw data, reducing the need for manual intervention. Given sufficient data, DL models can achieve higher accuracy in complex tasks compared to traditional ML models. However, DL requires large amounts of labeled data to train effectively and significant computational power, often necessitating specialized hardware like graphics processing units (GPUs). Additionally, DL models are often perceived as "black boxes" with low interpretability, making it challenging for dental professionals to understand the decision-making process [7, 10, 11].

The key differences between ML and DL in dentistry usage stem from their complexity and data needs. While ML can be applied to more straightforward tasks with smaller datasets, DL is suited for more complex tasks but requires vast amounts of data and computational power. DL generally provides higher accuracy in tasks such as image analysis, but ML models are more interpretable and easier to integrate into clinical workflows. ML's requirement for manual feature selection can be both a limitation and an advantage, offering better control over the process. In contrast, DL automates feature extraction but at the cost of increased complexity and resource requirements.

In summary, machine and deep learning have significant roles in dentistry. The choice between them depends on the task's specific requirements, the available data, and the need for interpretability versus accuracy. The number of research publications, including dentistry and AI, has increased in recent years, from only 27 in 2019 to 414 in 2024. AI applications can be used in several areas, such as recognizing anatomic structures [12-14], oral rehabilitation [15], periodontology [16, 17], dental implants [18], endodontics [19, 20], and pathologies. However, most studies of AI in our field are related to radiology since it has the most data and can provide quantitative and qualitative data.

ADVANCING ORTHODONTIC CARE

In orthodontics, AI has become a valuable tool for enhancing diagnostic accuracy and treatment planning. Integrating AI technologies, such as ML and neural networks, allows for more precise analysis of dental images and patient data. This leads to better-informed decisions and more effective orthodontic treatments. With continuous advancements, AI is revolutionizing orthodontic care, making it more efficient and personalized for each patient. Diagnosis and treatment planning are the primary and probably most important elements of a treatment. AI can assist orthodontists in identifying the patient's cervical vertebrae maturation with high accuracy and can also compare it to the current diagnostic standard (i.e., hand-wrist radiographs) [21-23]. Another usage is for making orthodontic extraction decisions, where AI had a 93% success rate for determining the need of extraction, 89% in choosing which tooth should be extracted, and 96% in defining if anterior retraction will be needed after extraction [24]. The potential risk for obstructive sleep apnea can also be identified using lateral radiographs with high-performance metrics using DL [25]. Integrating AI models can detect temporomandibular joint osteoarthritis, mainly when integrating images with clinical examination and biomolecular aspects [26, 27]. AI can also assist orthodontists in making decisions regarding orthognathic surgery [28, 29].

DECISION MAKING PROCESS: CLASSIFICATION SYSTEMS AND TREATMENT STRATEGIES FOR MALOCCLUSION IN ORTHODONTICS AND ORTHOGNATHIC SURGERY

In orthodontics, various classification systems for malocclusions have been developed and studied to standardize diagnosis and facilitate communication among professionals. The widely used 1899 Angle classification is based on the anteroposterior relationship of the first permanent molars into Class I, Class II, and Class III, with their divisions and subdivisions [30]. In 1911, Lischer complemented this classification by considering the individual positions of each tooth within the arch [31]. In 1974, Andrews published a classification to determine the occlusal relationship more broadly, called the six keys to normal occlusion, where the first key includes the modified Angle classification. In his book, *Straight Wire: The Concept and Appliance*, malocclusions were divided into normal and abnormal, with normal malocclusion being treatable by orthodontics, while abnormal malocclusion requires the assistance of other specialties, such as orthognathic surgery [32].

In 2006, Reis researched facial analysis through the Aesthetic Pleasantness Pyramid and classified individuals as aesthetically pleasing, acceptable, or unpleasant, to facilitate communication with patients and laypeople. According to the article, the orthodontist's role is to keep patients between pleasant and acceptable boundaries. For that, patients classified as unpleasant should undergo orthodontic-surgical procedures [33]. In 2004, Capelozza used craniofacial growth pattern as the main etiological factor and concluded that when genetic evolution has a negative tendency to impact both occlusion and the patient's face, orthognathic surgery should be performed with orthodontic treatment [34]. Despite various ways to classify malocclusion, it is agreed that individuals with significant skeletal discrepancies that prevent compensatory orthodontic treatment should undergo orthodontic-surgical treatments to correct occlusion, aesthetics, and function. However, quantitative analyses are necessary to classify the need for orthodontic or surgical treatment. The literature presents quantitative and qualitative methods to evaluate discrepancies through radiographs, models, soft tissue analysis, computed tomography, and subjective facial analyses [33, 35-38]. These methods help the orthodontist better understand the location, etiology, and severity of the patient's problem. These elements, along with the patient's age and facial attractiveness, are crucial for the orthodontist to outline the prognosis and treatment plan.

The patient's treatment can be corrective orthodontics for patients with balanced bone bases and face; compensatory orthodontics for patients with acceptable skeletal and facial discrepancies; or decompensatory surgical orthodontics for patients with significant skeletal discrepancies and facial unpleasantness that prevent orthodontic treatment alone from restoring aesthetics, occlusion, and function [34, 39-44]. A significant challenge in compensatory treatment is borderline patients who present with two treatment possibilities: dental compensation, which consists of axially inclining the teeth in the opposite direction to the patient's skeletal error, or treatment with orthodontic decompensation and orthognathic surgery. In these cases, the decision-making will depend on the orthodontist's clinical experience, quantitative and qualitative scientific analyses, and the patient's perception of the problem related to self-esteem and social issues [28, 44, 45]. Figure 1 shows an example of workflow for treatment decisions.

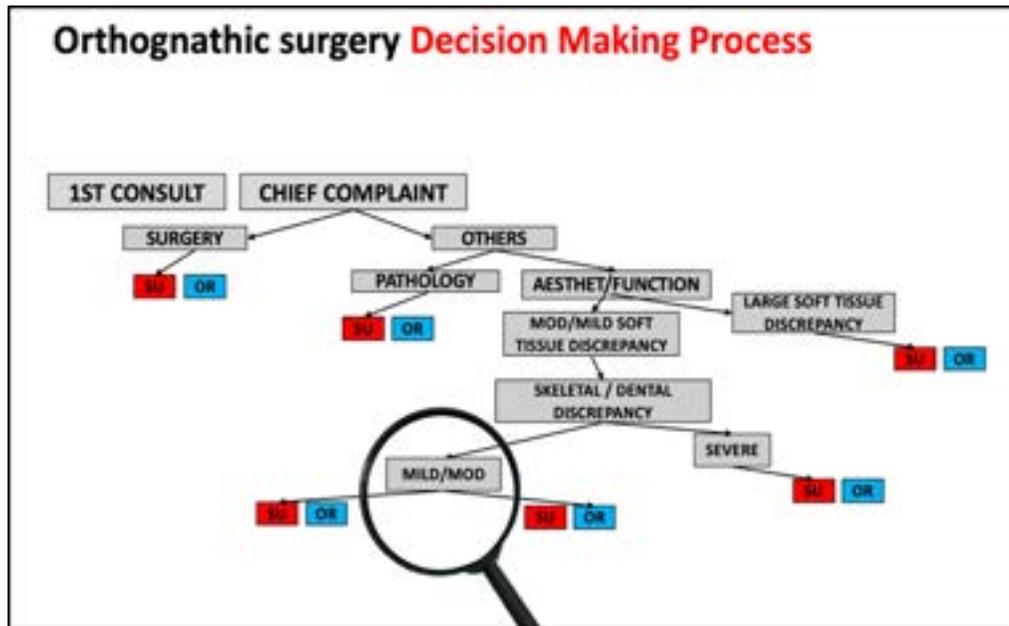


Figure 1. Decision-making process workflow for orthognathic surgery in borderline patients, considering dental compensation and orthognathic surgery options based on clinical assessment and patient factors. Abbreviations: AESTHET, aesthetic; MOD, moderate; OR, orthodontics; SU, surgery.

In 1959, Steiner demonstrated the ideal positioning of incisors in orthodontically treated patients to achieve a "pleasant face." This positioning is based on the ANB angle, the inclination of the upper and lower incisors, and their distances from the NA and NB lines. He also clarified that, depending on the type of malocclusion, these normal values might need to be adjusted to achieve satisfactory results. Additionally, Steiner provided estimates considered ideal for orthodontic compensation in patients with skeletal discrepancies, aiming to establish limits for dental movements. While acknowledging the existence of more advanced methods for guiding treatment than fixed normality standards and compensation limits, Steiner emphasized that these graphs could serve as valuable decision-making aids for less experienced professionals [46].

Other studies have attempted to quantify the need for orthognathic surgery. With new technologies such as skeletal anchorage and tomography, orthodontic limits have expanded, allowing for greater compensations. A classic study in this decision-making process is Proffit's envelope of discrepancy. However, isolated measurements are generally insufficient to define a correct final diagnosis. This is why some authors have described cephalometric correlations as essential aids in planning and decision-making. Subjective judgment and empirical knowledge are still necessary to determine whether a patient requires orthognathic surgery [28, 47, 48].

ARTIFICIAL INTELLIGENCE IN ORTHOGNATHIC SURGERY: ENHANCING DECISION-MAKING THROUGH CEPHALOMETRIC ANALYSIS AND MACHINE LEARNING

In the field of orthognathic surgery, AI has significantly advanced various aspects of the process. It has enhanced diagnosis and treatment planning, as well as improved patient expectations and

comprehension of the surgical procedure. Several studies have explored the use of AI to better understand the primary concerns patients have regarding orthognathic surgery and how AI can assist them in making informed decisions [49, 50]. Multiple researchers have employed AI to determine whether a patient requires orthognathic surgery, with the best results observed in Class III patients exhibiting different anteroposterior skeletal relationships. This likely occurs because the limitations for these cases are more pronounced, making compensations less acceptable. Most of these studies utilized supervised ML, where the algorithm learns from pre-labeled data based on the actual treatment received by the patient. This approach can be influenced by biases, such as the clinician's experience level or the patient's concerns [28, 51-53]. This probably happens because the limitations for these cases are higher, and compensations are less acceptable [54]. The majority of these studies rely on cephalometric radiographs, as cephalometric variables enable the use of ML models to identify and highlight the most critical measures for predicting the need for orthognathic surgery. In Figure 2, we present using J48, a decision tree algorithm, to classify the most important features in Class III patients, with 10-fold cross validation, while Figure 3 shows the same classification performance for Class II patients.

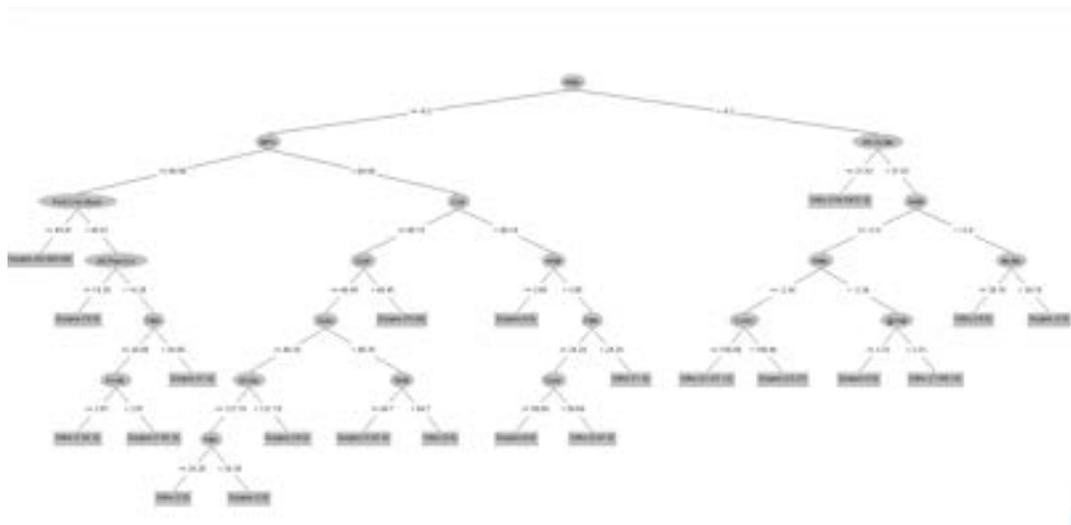


Figure 2. Classification of key features in Class III patients using J48 Decision Tree Algorithm with 10-fold cross validation. Abbreviations: IMPA, Incisor Mandibular Plane Angle; Wits, Wits Appraisal; SN.GoMe, Sella-Nasion to Gonion-Menton Angle; CoA, Condylion to Point A; GoAr, Gonion to Articulare; PostCran.Base, Posterior Cranial Base; SN.PlanOccl, Sella-Nasion to Occlusal Plane; FMA, Frankfort Mandibular Plane Angle; FH.Mx, Frankfort Horizontal to Maxilla; SN.Ba, Sella-Nasion to Basion; SNB, Sella-Nasion to Point B; HNB, Hyoid to Nasion-B; CoGn, Condylion to Gnathion; Mx.Md, Maxilla-Mandible Relationship; @1NB, Distance from the lower incisor to Nasion-B; FH, Frankfort Horizontal; Ortho, Orthodontic Treatment; Surgery, Surgical Treatment.

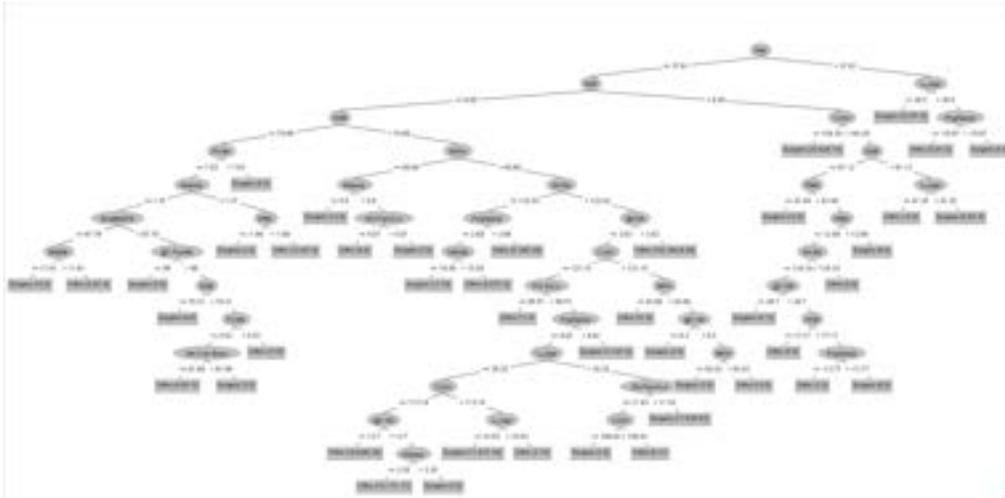


Figure 3. Classification of key features in Class II patients using J48 Decision Tree Algorithm with 10-fold cross validation. Abbreviations: SNB, Sella-Nasion to Point B; ANB, A Point - Nasion - B Point; FMA, Frankfort Mandibular Plane Angle; CoGn, Condylion to Gnathion; PogNperp, Pogonion to Nasion Perpendicular; ANperp, A Point to Nasion Perpendicular; SN.Ba, Sella-Nasion to Basion; SN.GoMe, Sella-Nasion to Gonion-Menton Angle; SN.PlanOccl, Sella-Nasion to Occlusal Plane; FMIA, Frankfort Mandibular Incisor Angle; Witts, Wits Appraisal; @1_NA, Upper Incisor to Nasion-A; @1_NB, Lower Incisor to Nasion-B; IMPA, Incisor Mandibular Plane Angle; @1.PlanMx, Upper Incisor to Maxillary Plane; Prm.SnLs, Primate Space at Subnasale Level; Ant.Cran.Base, Anterior Cranial Base; BfaBfp, Basion-Facial Axis to Basion-Facial Plane; NfaNfp, Nasion-Facial Axis to Nasion-Facial Plane; i_Age, Indexed Age; EnaMeAFAL, Lower Anterior Facial Height (LAFH); Ortho, Orthodontic Treatment; Surgery, Surgical Treatment.

Our results also have shown that the Wits appraisal was the most significant feature in the decision-making process, followed by vertical measurements and the lower incisor angle (Figure 4). Subsequent research confirmed that Wits and ANB were the most important features, with the occlusal plane angle also being highly significant [51, 55-58].

ENHANCING ORTHOGNATHIC SURGERY DECISION-MAKING: THE ROLE OF CLINICAL JUDGMENT AND DEEP LEARNING

Clinical judgment plays a critical role in the decision-making process for orthodontic treatment and orthognathic surgery. Figure 5 illustrates the number of recommendations for orthodontic treatment versus orthognathic surgery made by two surgeons and two orthodontists, based on a pool of approximately 1,000 patients. This figure highlights the level of agreement among these professionals. Despite their expertise, a significant degree of non-concordance is observed when comparing their recommendations. This discrepancy underscores the importance of clinical judgment and the variability inherent in professional assessments. Even among highly trained specialists, differing opinions can arise, emphasizing the need for comprehensive evaluation and individualized treatment planning.

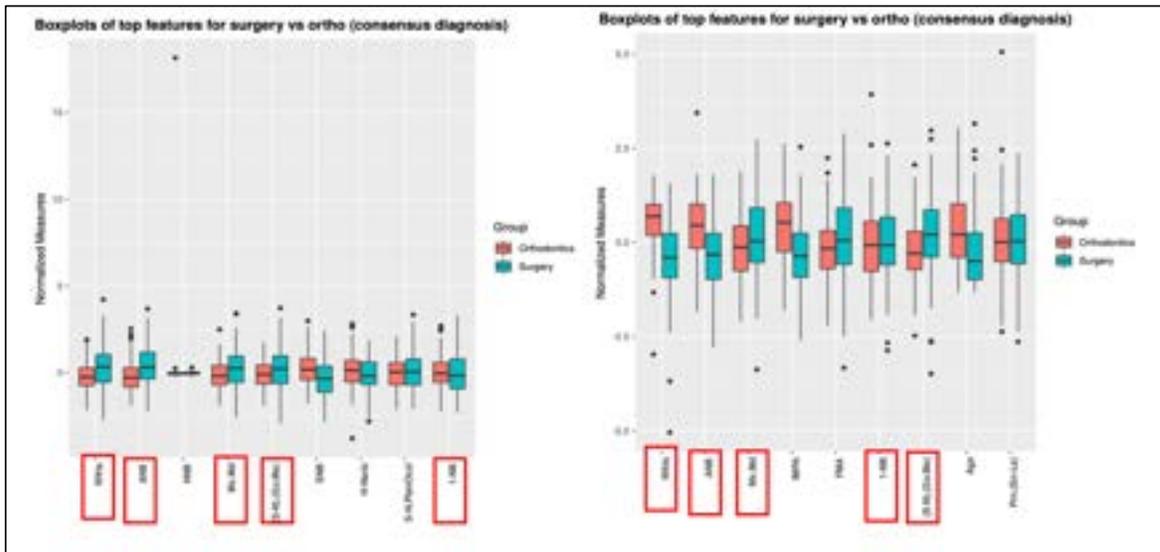


Figure 4. Boxplots of top features for consensus diagnosis in Class II (A) and Class III (B) patients: comparing surgery vs. orthodontic treatment.

Abbreviations: Wits, Wits Appraisal; ANB, A Point - Nasion - B Point; HNB, Profile's evaluation; Mx.Md, Maxillary to mandibular plane angle; SN.GoMe, Sella-Nasion to Gonion-Menton Angle; SNB, Sella-Nasion to Point B; FH.Mx, Frankfort to maxillary plane angle; S-N.PlanoOccl, Sella-Nasion to Occlusal Plane Angle; IMPA, Incisor Mandibular Plane Angle; FMA, Frankfort Mandibular Plane Angle; Age, Indexed Age; Prn.(Sn-Ls), Nasolabial angle.

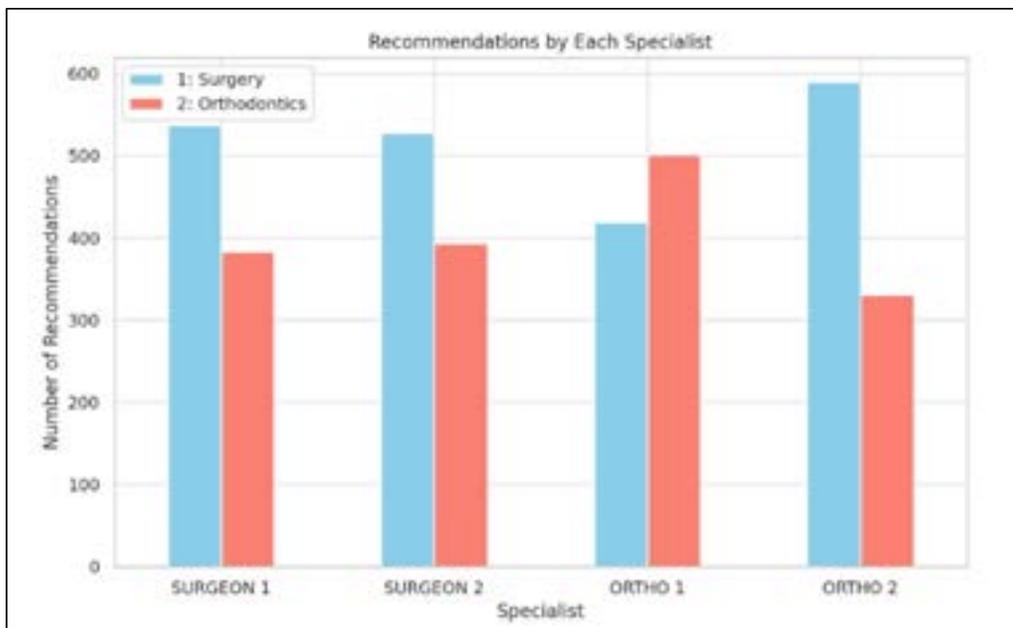


Figure 5. Recommendations for orthodontic treatment vs. orthognathic surgery by two surgeons and two orthodontists.

Figure 6 further explores this by showing the agreement rates among the specialists. The agreement rate between the two surgeons, the two orthodontists, and among all four specialists is presented. This data reveals that while there is some level of agreement, notable differences remain, reinforcing the necessity of integrating clinical expertise with advanced diagnostic tools, such as AI, to enhance consistency and accuracy in treatment decisions.

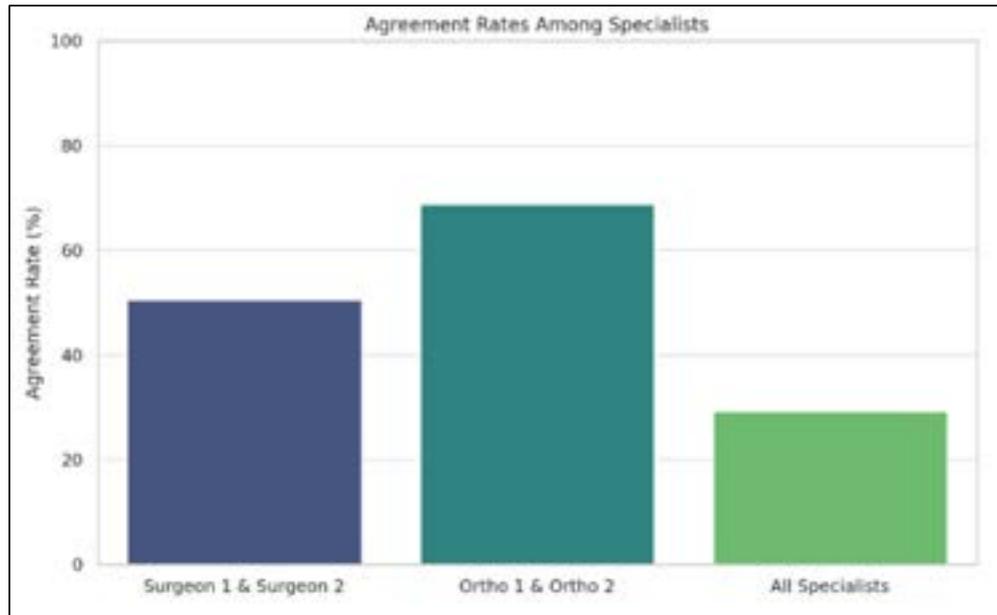


Figure 6. Agreement rates among specialists for orthodontic treatment vs. orthognathic surgery decisions.

DL can also be utilized to assist in orthognathic surgery. One of the key applications of DL is the analysis of raw images to create patterns and recognize features. This enables the use of patients' photographs or radiographic images, without any accompanying reports or numerical variables, to identify critical regions that directly impact the decision to refer a patient for orthognathic surgery. Furthermore, DL can aid in predicting or classifying skeletal patterns and assessing facial attractiveness, thereby enhancing 3D surgical simulations. Given that most studies have relied on lateral radiographs or photographs, future research should focus on 3D evaluations of patients. Additionally, it is crucial to include clinical reports and other subjective data in the AI training sets to ensure that the system can comprehend and address the needs and expectations of patients effectively [59-61]. In summary, while clinical judgment remains essential in treatment planning, the integration of DL and other AI technologies offers significant potential to improve decision-making in orthognathic surgery. By combining expert clinical insights with advanced data analysis tools, the field can move towards more precise, consistent, and personalized patient care.

CONCLUSIONS

The integration of AI in orthognathic surgery has significantly enhanced various aspects of patient care, including diagnosis, treatment planning, and patient education. AI technologies, particularly ML and DL, have demonstrated their potential to improve diagnostic accuracy and support clinical decision-making by analyzing extensive datasets and identifying critical cephalometric variables. The application of AI in determining the need for orthognathic surgery has shown promising results, especially in Class III patients with distinct anteroposterior skeletal relationships. The use of supervised ML models has enabled the identification of key features, such as the Wits appraisal, ANB angle, and

occlusal plane angle, which are crucial in the decision-making process. These models, while beneficial, must be carefully managed to mitigate biases arising from subjective clinical judgments and patient-specific factors. DL, with its ability to analyze complex data types and recognize patterns in raw images, offers additional capabilities for AI in orthognathic surgery. DL models can assess patients' photographs and radiographs, aiding in the prediction and classification of skeletal patterns and facial attractiveness. This enhances 3D surgical simulations and provides a more comprehensive understanding of patient needs and expectations. Despite the advancements, the use of AI in orthognathic surgery requires further exploration, particularly in incorporating 3D evaluations and subjective clinical data into AI training sets. This will ensure that AI systems can better understand and address patients' expectations and clinical needs. In summary, AI is not intended to replace human expertise but to augment it, providing orthodontists and maxillofacial surgeons with advanced tools to support high-quality diagnostics and treatment plans. The continuous advancements in AI technologies promise to further revolutionize orthognathic surgery, making it more precise, efficient, and patient-centered.

REFERENCES

1. Kudyba SP. *Healthcare Informatics: Improving Efficiency and Productivity, 1st Ed.* CRC Press 2010.
2. Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. *Health Inf Sci Syst* 2014;2:3.
3. Leite AF, Vasconcelos KF, Willems H, Jacobs R. Radiomics and machine learning in oral healthcare. *Proteomics Clin Appl* 2020;14(3):e1900040.
4. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, et al. Artificial intelligence in healthcare: past, present and future. *Stroke Vasc Neurol* 2017;2(4):230-243.
5. Schwendicke F, Chaurasia A, Arsiwala L, Lee J-H, Elhennawy K, Jost-Brinkmann PG, et al. Deep learning for cephalometric landmark detection: systematic review and meta-analysis. *Clin Oral Investig* 2021;25(7):4299-4309.
6. Machoy ME, Szyzka-Sommerfeld L, Vegh A, Gedrange T, Woźniak K. The ways of using machine learning in dentistry. *Adv Clin Exp Med* 2020;29(3):375-384.
7. Corbella S, Srinivas S, Cabitza F. Applications of deep learning in dentistry. *Oral Surg Oral Med Oral Pathol Oral Radiol* 2021;132(2):225-238.
8. Mehrabani M. Artificial intelligence in dentistry. *Br Dent J.* 2024;236:522–522.
9. Ghahramani Z. Probabilistic machine learning and artificial intelligence. *Nature* 2015; 521(7553):452-459.
10. Kim H, Shim E, Park J, Kim YJ, Lee U, Kim Y. Web-based fully automated cephalometric analysis by deep learning. *Comput Methods Programs Biomed* 2020;194:105513.
11. Seo H, Hwang J, Jeong T, Shin J. Comparison of deep learning models for cervical vertebral maturation stage classification on lateral cephalometric radiographs. *J Clin Med* 2021;10 (16):3591.
12. Miki Y, Muramatsu C, Hayashi T, Zhou X, Hara T, Katsumata A, et al. Classification of teeth in cone-beam CT using deep convolutional neural network. *Comput Biol Med* 2017;80:24-29.
13. Chen H, Zhang K, Lyu P, Li H, Zhang L, Wu J, et al. A deep learning approach to automatic teeth detection and numbering based on object detection in dental periapical films. *Sci Rep* 2019;9 (1):3840.
14. Vinayahalingam S, Xi T, Bergé S, Maal T, de Jong G. Automated detection of third molars and mandibular nerve by deep learning. *Sci Rep* 2019;9(1):9007.
15. Devito KL, de Souza Barbosa F, Filho WN. An artificial multilayer perceptron neural network for diagnosis of proximal dental caries. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 2008;106(6):879-884.

16. Lee JH, Kim DH, Jeong SN, Choi SH. Diagnosis and prediction of periodontally compromised teeth using a deep learning-based convolutional neural network algorithm. *J Periodontal Implant Sci* 2018;48(2):114-123.
17. Krois J, Ekert T, Meinhold L, Golla T, Kharbot B, Wittemeier A, et al. Deep learning for the radiographic detection of periodontal bone loss. *Sci Rep* 2019;9(1)8495.
18. Lee JH, Kim YT, Lee JB, Jeong SN. A performance comparison between automated deep learning and dental professionals in classification of dental implant systems from dental imaging: a multi-center study. *Diagnostics (Basel)* 2020;10(11):910.
19. Saghiri MA, Asgar K, Boukani KK, Lotfi M, Aghili H, Delvarani A, et al. A new approach for locating the minor apical foramen using an artificial neural network. *Int Endod J* 2012;45(3):257-265.
20. Kositbowornchai S, Plermkamon S, Tangkosol T. Performance of an artificial neural network for vertical root fracture detection: an ex vivo study. *Dent Traumatol* 2013;29(2):151-155.
21. Kök H, İzgi MS, Acilar AM. Determination of growth and development periods in orthodontics with artificial neural network. *Orthod Craniofac Res* 2021;24 Suppl 2:76-83.
22. Kök H, Acilar AM, İzgi MS. Usage and comparison of artificial intelligence algorithms for determination of growth and development by cervical vertebrae stages in orthodontics. *Prog Orthod* 2019;20(1):41.
23. Kim DW, Kim J, Kim T, Kim T, Kim YJ, Song IS, et al. Prediction of hand-wrist maturation stages based on cervical vertebrae images using artificial intelligence. *Orthod Craniofac Res* 2021;24(Suppl 2):68-75.
24. Takada K. Artificial intelligence expert systems with neural network machine learning may assist decision-making for extractions in orthodontic treatment planning. *J Evid Based Dent Pract* 2016;16(3):190-192.
25. Tsuiki S, Nagaoka T, Fukuda T, Sakamoto Y, Almeida FR, Nakayama H, et al. Machine learning for image-based detection of patients with obstructive sleep apnea: an exploratory study. *Sleep Breath* 2021;25(4):2297-2305.
26. Bianchi J, de Oliveira Ruellas AC, Gonçalves JR, Paniagua B, Prieto JC, Styner M, et al. Osteoarthritis of the temporomandibular joint can be diagnosed earlier using biomarkers and machine learning. *Sci Rep* 2020;10(1):8012.
27. Lee KS, Kwak HJ, Oh JM, Jha N, Kim YJ, Kim W, et al. Automated detection of TMJ osteoarthritis based on artificial intelligence. *J Dent Res* 2020;99(12):1363-1367.
28. de Oliveira PHJ, Li T, Li H, Gonçalves JR, Santos-Pinto A, Gandini Junior LG, et al. Artificial intelligence as a prediction tool for orthognathic surgery assessment. *Orthod Craniofac Res* 2024; 27(5):785-794.
29. Choi H Il, Jung SK, Baek SH, Lim WH, Ahn SJ, Yang IH, et al. Artificial intelligent model with neural network machine learning for the diagnosis of orthognathic surgery. *J Craniofac Surg* 2019;30(7):1986-1989.
30. Angle EH. Classification of malocclusion. *Dental Costomos* 1899;41(3):248-264.
31. Lischer BE. *Principles and Methods of Orthodontics: An Introductory Study of the Art for Students and Practitioners of Dentistry*. Lea & Febiger 1912.
32. Andrews LF. *Straight Wire: The Concept and Appliance*. San Diego, CA: L.A. Wells Co. 1989.
33. Reis SAB, Abrão J, Capelozza Filho L, Claro CA de A. Análise Facial Subjetiva. *Revista Dental Press de Ortodontia e Ortopedia Facial*. 2006;11(5):159-172.
34. Capelozza Filho L. *Diagnóstico em Ortodontia*. Maringá, PR: Dental Press 2004.
35. Arnett GW, Jelic JS, Kim J, Cummings DR, Beress A, Worley CM, et al. Soft tissue cephalometric analysis: diagnosis and treatment planning of dentofacial deformity. *Am J Orthod Dentofacial Orthop* 1999;116:239-253.
36. de Oliveira MA, Pereira MD, Hino CT, Campaner AB, Scanavini MA, Ferreira LM. Prediction of transverse maxillary dimension using orthodontic models. *J Craniofac Surg* 2008;19(6):1465-1471.

37. Miller AJ, Maki K, Hatcher DC. New diagnostic tools in orthodontics. *Am J Orthod Dentofacial Orthop* 2004;126(4):395-396.
38. Allareddy V, Rengasamy Venugopalan S, Nalliah RP, Caplin JL, Lee MK, Allareddy V. Orthodontics in the era of big data analytics. *Orthod Craniofac Res* 2019;22(Suppl 1):8-13.
39. Wolford LM. Orthodontics for orthognathic surgery. In: Miloro, M, Ghali, GE, Larsen, PE, Waite, P (eds). *Peterson's Principles of Oral and Maxillofacial Surgery*. Springer, Cham 2022.
40. Johnston C, Hunt O, Burden D, Stevenson M, Hepper P. Self-perception of dentofacial attractiveness among patients requiring orthognathic surgery. *Angle Orthod* 2010;80(2):361-366.
41. Gerzanic L, Jagsch R, Watzke IM. Psychologic implications of orthognathic surgery in patients with skeletal Class II or Class III malocclusion. *Int J Adult Orthodon Orthognath Surg* 2002;17(2):75-81.
42. Naini FB, Gill DS, Eds. *Orthognathic Surgery: Principles, Planning and Practice*. John Wiley & Sons Ltd 2017.
43. Garvill J, Garvill H, Kahnberg KE, Lundgren S. Psychological factors in orthognathic surgery. *J Craniomaxillofac Surg* 1992;20(1):28-33.
44. Raposo R, Peleteiro B, Paço M, Pinho T. Orthodontic camouflage versus orthodontic-orthognathic surgical treatment in Class II malocclusion: a systematic review and meta-analysis. *Int J Oral Maxillofac Surg*. 2018;47(4):445-55.
45. van Steenbergen E, Litt MD, Nanda R. Presurgical satisfaction with facial appearance in orthognathic surgery patients. *Am J Orthod Dentofacial Orthop* 1996;109(6):653-659.
46. Steiner CC. Cephalometrics in clinical practice. *Angle Orthod* 1959;29(1):8-29.
47. Proffit WR, Fields JR HW, Sarver DM, Larson B. *Contemporary Orthodontics, 6th ed*. Elsevier 2018.
48. Graber L, Vanarsdall R, Vig K, Huang G. *Orthodontics: Current Principles and Techniques, 6th ed*. Elsevier 2017.
49. Fatima K, Singh P, Amipara H, Chaudhary G. Accuracy of artificial intelligence-based virtual assistants in responding to frequently asked questions related to orthognathic surgery. *J Oral Maxillofac Surg* 2024.82(8):916-921.
50. Salazar D, Rossouw PE, Javed F, Michelogiannakis D. Artificial intelligence for treatment planning and soft tissue outcome prediction of orthognathic treatment: a systematic review. *J Orthod* 2024;51(2):107-119.
51. Shin W, Yeom HG, Lee GH, Yun JP, Jeong SH, Lee JH, et al. Deep learning based prediction of necessity for orthognathic surgery of skeletal malocclusion using cephalogram in Korean individuals. *BMC Oral Health* 2021;21(1):130.
52. Taraji S, Atici SF, Viana G, Kusnoto B, Allareddy V (Sath), Miloro M, et al. Novel machine learning algorithms for prediction of treatment decisions in adult patients with Class III malocclusion. *J Oral Maxillofac Surg* 2023;81(11):1391-1402.
53. Lee H, Ahmad S, Frazier M, Dundar MM, Turkkahraman H. A novel machine learning model for Class III surgery decision. *J Orofac Orthop* 2024;85(4):239-249.
54. Kerr WJ, Miller S, Dawber JE. Class III malocclusion: surgery or orthodontics? *Br J Orthod* 1992;19(1):21-24.
55. Lin HH, Chiang WC, Yang CT, Cheng CT, Zhang T, Lo LJ. On construction of transfer learning for facial symmetry assessment before and after orthognathic surgery. *Comput Methods Programs Biomed* 2021;200:105928.
56. Tanikawa C, Yamashiro T. Development of novel artificial intelligence systems to predict facial morphology after orthognathic surgery and orthodontic treatment in Japanese patients. *Sci Rep* 2021;11(1):15853.
57. de Oliveira PHJ, Li T, Li H, Gonçalves JR, Santos-Pinto A, Gandini Junior LG, et al. Artificial intelligence as a prediction tool for orthognathic surgery assessment. *Orthod Craniofac Res* 2024;27(5):785-794.

58. Barone S, Cevidanes L, Miranda F, Gurgel ML, Anchling L, Hutin N, et al. Enhancing skeletal stability and Class III correction through active orthodontist engagement in virtual surgical planning: a voxel-based 3-dimensional analysis. *Am J Orthod Dentofacial Orthop* 2024;165(3):321-331.
59. Lee KS, Ryu JJ, Jang HS, Lee DY, Jung SK. Deep convolutional neural networks based analysis of cephalometric radiographs for differential diagnosis of orthognathic surgery indications. *App Sci* 2020;10(6):2124.
60. Yu HJ, Cho SR, Kim MJ, Kim WH, Kim JW, Choi J. Automated skeletal classification with lateral cephalometry based on artificial intelligence. *J Dent Res* 2020;99(3):249-256.
61. Patcas R, Bernini DAJ, Volokitin A, Agustsson E, Rothe R, Timofte R. Applying artificial intelligence to assess the impact of orthognathic treatment on facial attractiveness and estimated age. *Int J Oral Maxillofac Surg* 2019;48(1):77-83.

REVIEW ARTICLE

Leveraging Generative Artificial Intelligence in Teaching, Scholarship and Dental Education: Use Cases and Reflections

Jonas Bianchi¹  | Meixun Zheng² ¹Department of Orthodontics, Arthur A. Dugoni School of Dentistry, University of the Pacific, San Francisco, California, USA | ²Office of Academic Affairs, Arthur A. Dugoni School of Dentistry, University of the Pacific, San Francisco, California, USA**Correspondence:** Jonas Bianchi (jbianchi@pacific.edu)**Received:** 3 February 2025 | **Revised:** 18 April 2025 | **Accepted:** 30 May 2025**Funding:** This work was supported by American Association of Orthodontists Foundation, Robert E. Binder Teaching Fellowship Award.**Keywords:** artificial intelligence in dentistry | dental education | generative AI | machine learning

ABSTRACT

The objective of this narrative review is to explore the role of generative artificial intelligence (genAI) in dental education, highlighting its emerging applications, potential benefits and implementation challenges. Since the launch of ChatGPT in 2022, genAI tools have gained traction in academic and clinical settings, enabling content generation, translation, summarisation, exam preparation and basic clinical planning. This review presents a series of illustrative use cases demonstrating how genAI has been integrated into teaching, research and clinical workflows in dental and orthodontic training. Each example underscores how AI can support faculty in course design, assist students with learning complex concepts and provide real time feedback for exam analysis and academic writing. However, the implementation of genAI is not without limitations. The review addresses common concerns, including misinformation, data privacy, fabricated references and ethical use in clinical contexts. Although the benefits of genAI are promising, this review emphasises the importance of human oversight and institutional policies to guide ethical and effective use. In conclusion, genAI offers valuable support in dental education when used responsibly. Continued dialogue among educators, students and policymakers is essential to ensure that AI tools are integrated thoughtfully and equitably into academic practice.

1 | Introduction

Artificial intelligence (AI) has become integral to our daily lives. According to Google, AI is 'a field of science concerned with building computers and machines that can reason, learn, and act in such a way that would normally require human intelligence or that involves data whose scale exceeds what humans can analyze' [1]. As the field of AI has expanded, the definition of AI has also become more multifaceted, encompassing a variety of techniques that allow machines to perform various tasks that were originally only doable by human beings. A new type of AI, generative AI (hereafter referred

to as 'genAI'), enables machines to create new content, including text, audio, music, images, videos and codes, in response to user prompts [2]. Since OpenAI released ChatGPT in November 2022, genAI has become ubiquitous across various sectors, including education. ChatGPT can engage in both text and audio conversations, write essays and summaries, generate codes in different programming languages, perform advanced data analysis and visualisations and solve complex reasoning tasks, such as mathematical problem-solving. Alongside the development of ChatGPT, other genAI tools have emerged, including Microsoft's Co-Pilot, Google's Bard and Anthropic's Claude AI. Additionally, Google introduced

Med-PaLM, a genAI model designed for the medical field and the first large language model to pass the U.S. Medical Licensing Exam (USMLE) [3]. More recently, search engines like Google have integrated the 'AI overview' feature, which provides summarised responses to user queries, helping users navigate the vast amount of internet information more efficiently. Researchers across various disciplines have explored both the opportunities and limitations of genAI in education [4–10]. A recent literature review [11] concluded that genAI offers several exciting opportunities in medical education, including promoting self-directed learning, creating simulated scenarios and assisting with writing. Similarly, a scoping review by Furey et al. [12] examined the role of AI in medical education, particularly in emerging care. Although the use of AI in education has varied, genAI presents unique advantages, such as streamlining curriculum development, improving assessment methods, enhancing academic support for students and creating highly realistic clinical simulations. The authors argued that incorporating genAI into education can strengthen students' digital literacy and meet their diverse learning needs. In better the field of dental education, Ali et al. [13] investigated the potential of ChatGPT for assessment. This study evaluated ChatGPT's performance across different question types, including multiple-choice, short-answer and essay questions. Although ChatGPT demonstrated proficiency in knowledge-based assessments, it struggled with tasks requiring critical analysis of research and image interpretation. The study highlighted the need for educators to adapt their teaching and evaluation methods to integrate AI while maintaining academic integrity effectively. Rather than viewing genAI as a threat, the researchers suggested embracing it as a tool to enhance virtual learning and create more personalised educational experiences for students.

Recent studies have also explored the potential of chatbots in supporting clinical diagnosis and patient education. For instance, AI-driven conversational agents have been tested for triage support, dermatologic assessment and initial diagnostic guidance in outpatient settings [14]. In parallel, chatbots are being deployed to assist patients in understanding treatment options, managing chronic conditions, and navigating emotional or mental health concerns through structured dialogues. These developments highlight the broader utility of genAI beyond academic support and suggest opportunities for integration into dental patient communication and decision-making [15–17].

Although there is a lot of excitement around its potential, there is also widespread uncertainty and anxiety regarding its application in educational settings. Initial concerns have been centered around misinformation, copyright, data privacy and academic integrity [10, 18, 19]. Yet, as technology evolves, the overall sentiment has become more optimistic [20]. It has gradually shifted from banning its use to exploring ways to guide effective and responsible use of this technology [21]. Many institutions have responded by developing guidelines and policies that support cautious integration, exploration, and ongoing evaluation of genAI [22].

A 2023 Educause national poll found that 83% of higher education professionals believe genAI will significantly impact the higher education landscape within the next 3–5 years and that

its benefits outweigh its drawbacks [20]. Meanwhile, a 2024 national survey revealed that 77% of administrators and faculty at 4-year institutions feel unprepared for the changes AI is expected to bring. This discrepancy underscores a gap between the anticipated impact of genAI and institutions' readiness to adopt it effectively and maximise its benefits [23].

In response to this evolving landscape in higher education, peer-based use cases, collaborative efforts, and communities of practice are critical [20]. As we navigate this new technology, small-scale experiments can serve as valuable starting points. These efforts can help identify where genAI tools offer the greatest advantages in education. Reflecting on these applications is equally important, as it enables us to identify optimal uses while mitigating potential drawbacks. However, despite the rapidly growing body of research on genAI, a gap remains between theoretical discussions and practical implementation. This paper aims to help bridge this gap by focusing on how dental educators could utilise genAI to augment daily work. By presenting ideas and tools that span teaching and learning, educational research, and clinical practice, we demonstrate the diverse applications of genAI, reflect on the success and challenges, including data privacy and security management, and offer solutions to address these issues. We hope to inspire and support colleagues as they begin experimenting with this innovative technology in their practice.

In the following sections, we present a range of use cases illustrating the application of genAI tools, using ChatGPT as an example, in teaching, learning and research. For each case, we provide the prompts used, the AI output, and a brief evaluation and reflection. Additionally, we discuss challenges and provide tips to mitigate potential risks.

2 | GENAI for Oral Health-Related Information

When discussing genAI, many immediately associate it with student cheating. Although this is a valid issue, it is essential to look beyond it. In fact, learning to use generative AI effectively and responsibly can be a valuable skill for students. Rather than serving as a shortcut, genAI can be used as a powerful learning tool.

2.1 | Use Case 1: Explain Dental Concepts in Plain Language

Prompt: *Explain the step-by-step process of an implant procedure to me in plain language.*

Output: GPT provided the step-by-step explanation of the procedure, as quoted below:

'Dental implants are a great solution for missing teeth, and explaining the process in simple terms can be very helpful for your patients. Here's a basic step-by-step breakdown in plain language:

1. Consultation and Planning: Before anything happens, you'll meet with your dentist or oral surgeon to discuss the

procedure, take X-rays or 3D images and plan out the best approach.

2. **Tooth Extraction:** If the tooth you're replacing is still in your mouth, it'll need to be removed first.
3. **Bone Grafting (if necessary):** Sometimes, if the jawbone isn't thick or strong enough, a small piece of bone is added to make sure the implant has a solid foundation. This could be from another part of your body, a donor or synthetic material.
4. **Implant Placement:** Here's where the magic happens! The dentist or oral surgeon drills a small hole into your jawbone and places the dental implant, which is basically a titanium screw that acts like the root of your tooth.
5. **Healing and Integration (Osseointegration):** This is the waiting game. Over several months, the implant will fuse with your bone, becoming a strong anchor for the new tooth. This is why implants are so sturdy.
6. **Abutment Placement:** Once the implant is secure, a small connector (called an abutment) is placed on top of it. This will hold the new tooth in place.
7. **Impression and Crown Creation:** Finally, a mould or scan is taken to create a custom crown that fits perfectly and matches the rest of your teeth. The crown is then attached to the abutment, completing your new smile!

2.1.1 | Evaluation

The response is accurate and well-structured, making it accessible to individuals without a dental background. GenAI tools like ChatGPT can assist students in explaining complex dental terminology to patients who may have little to no prior knowledge of dentistry. This function is particularly valuable for students who are non-native speakers, helping them communicate more effectively with patients. By leveraging AI-generated explanations, students can enhance patient education as well as their own communication skills.

2.2 | Use Case 2: Translating Dental Procedures

GenAI tools like ChatGPT can assist with real time translation, enabling students and practitioners to communicate effectively with patients who speak different languages. Building on the previous case, students can ask ChatGPT to translate the step-by-step process of a dental implant procedure into the language spoken by the patient.

Prompt: *Translate the step-by-step process of the implant procedure into Mandarin Chinese.*

Output: ChatGPT translated the steps of the procedure as quoted below.

'Of course! Here's the step-by-step dental implant procedure from the initial response translated into Mandarin Chinese:

1. **Consultation and Planning:** 咨询和计划在开始任何的手术之前,您将与您的牙医或口腔外科医生会面,进行X光或3D影像检查,并计划最适合的治疗方案。
2. **Tooth Extraction:** 拔牙如果您需要替换的牙齿仍在口中,首先需要将其拔除。
3. **Bone Grafting (if necessary):**骨移植(如需要)有时,如果颌骨不够厚或不够坚固,需要添加一小块骨质来确保植体有一个坚固的基础。
4. **Implant Placement:** 植体放置这是“手术”的重头戏!医生会在下颌骨钻一个小孔,然后插入一个类似牙根的钛制螺丝。
5. **Healing and Integration (Osseointegration):** 愈合和骨整合(骨结合):这是等待的阶段。几个月后,植入体会与骨头融合,成为新牙的稳定基础。
6. **Abutment Placement:**基台放置一旦植体牢固,一个小的连接器(称为“基台”)将被放置在植体上,用来固定新的牙冠。
7. **Impression and Crown Creation:** 印模和牙冠制作最后,医生会进行印模或扫描,以制作出与您牙齿完美匹配的定制牙冠。
8. **Crown Placement:** 牙冠安装最后,将定制牙冠固定到基台之上。'

2.2.1 | Evaluation

The translation provided by ChatGPT is accurate and clear. By utilising AI-powered translation, students can enhance patient understanding and deliver more inclusive care. In fact, patients may have already used ChatGPT or similar AI tools before their appointment to research their procedure in their native language. This could influence communication dynamics, as patients might arrive with AI-generated information or misconceptions. As a result, dental students and practitioners may need to adapt their communication strategies to not only explain procedures clearly but also ensure that patients receive accurate guidance.

3 | GENAI in Teaching

GenAI can also offer valuable applications for educators. The following cases illustrate how dental educators can leverage ChatGPT in teaching and course design.

3.1 | Use Case 3: Course Planning

When preparing a course syllabus, an essential step is developing a weekly agenda (course calendar) that includes weekly class meeting dates. Manually checking the calendar and listing out each date can be time-consuming, especially when ensuring that live class sessions do not overlap with national holidays. AI tools can streamline this process.

Prompt: *I teach a course that meets every Tuesday in the Spring semester of AY 2024–2025. Give me a list of all Tuesdays from Jan to May 2025 in this format: Month/Day. If a specific week falls on a holiday, add the holiday name next to the date.*

Output: ChatGPT generated the requested list of Tuesdays, including any holidays that may impact class scheduling.

‘Here is the list of all Tuesdays from January to May 2025, including holidays:

[“01/07,” “01/14,” “01/21 (Martin Luther King Jr. Day),” “01/28,” “02/04,” “02/11,” “02/18 (Presidents’ Day),” “02/25,” “03/04,” “03/11,” “03/18,” “03/25,” “04/01,” “04/08,” “04/15,” “04/22,” “04/29,” “05/06,” “05/13,” “05/20,” “05/27 (Memorial Day)”].’

3.1.1 | Evaluation

ChatGPT correctly listed all Tuesdays and noted national holidays. This capability can significantly reduce the time faculty spend on administrative tasks, allowing them to focus more on creative course design. However, it is strongly suggested that faculty always verify the accuracy of the response.

3.2 | Use Case 4: Exam Report Interpretation

Faculty can use ChatGPT to help interpret exam reports to gain insights into assessment quality and student performance. In the example below, we uploaded an aggregated exam performance report to evaluate how well ChatGPT could interpret key assessment metrics (Figure 1). Specifically, we tested its ability to interpret reliability metrics such as the KR-20 value, a widely used measure of internal consistency for multiple-choice exams.

Prompt: Look at the provided screenshot. Explain the data in this exam report.

Output: Below is the explanation provided by ChatGPT.

‘Assessment Score Reliability (KR-20): The Kuder-Richardson Formula 20 (KR-20) is a measure of internal consistency reliability for measures with dichotomous choices, like multiple-choice exams. Here, it’s 0.77, which falls in the “Good” range, suggesting

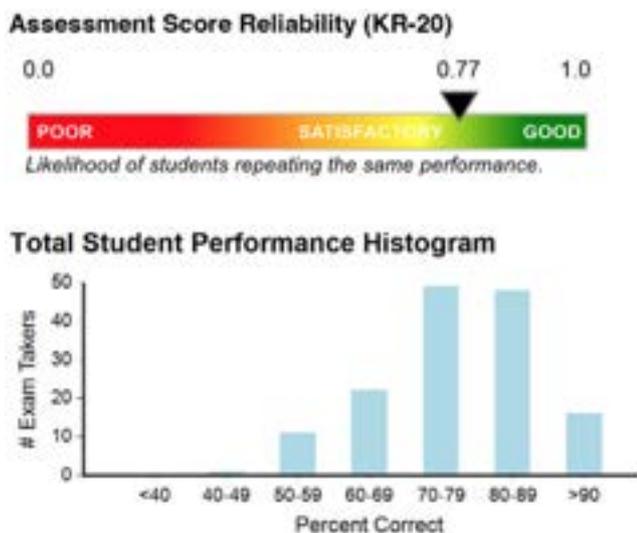


FIGURE 1 | Sample exam performance report.

that the exam has a good level of consistency and that you can be confident in the reliability of the exam results. This score indicates a moderate to high likelihood that students would perform consistently on repeated administrations of the test.

Total Student Performance Histogram: This is a graphical representation of the distribution of scores. It shows the number of students falling within certain percentage correct ranges. A quick look shows that the majority of students scored between 60%–69% and 70%–79%, with fewer students in the lower (< 40%) and higher (> 90%) brackets.’

3.2.1 | Evaluation

ChatGPT provided an accurate interpretation of the exam’s reliability metric, correctly explaining the significance of the KR-20 value in assessing the internal consistency of the exam. However, its interpretation of the histogram was partially incorrect. Although it correctly identified that many students scored between 60%–69% and 70%–79%, it overlooked students who scored between 80% and 89%. This example highlights an important limitation of genAI tools. Although AI can assist with data interpretation, it may generate incomplete or incorrect analyses. Therefore, users should always verify the accuracy of responses before drawing conclusions.

4 | GenAI In Educational Research

GenAI can support researchers throughout the entire research lifecycle, from conducting literature reviews to generating hypotheses, collecting and analysing data and presenting the findings. The following examples illustrate how ChatGPT can be used as a research aid, particularly in statistical analysis and visualisation.

4.1 | Use Case 5: Conducting Power Analysis

ChatGPT is a powerful coding tool that can assist with various statistical tests, including power analysis, to calculate the required sample size for an experiment. In the example below, we tested how well it can conduct power analysis.

Prompt: Calculate the sample size for a research study on [virtual reality]. I have an experimental group and a control group. I want the alpha level (p value) to be 0.05, the power to be 0.8, and the estimated effect size to be 0.4. The number of participants in each group will be the same. Adjust for 10% sample loss. Give me the number of participants I need for each group.

Output: ChatGPT wrote Python code to compute the sample size, with the calculation indicating 110 participants per group after adjusting for a 10% sample loss.

4.1.1 | Evaluation

A review of the code confirmed that the calculated sample size is correct, as seen in Figure 2. However, as demonstrated in this use case, the researcher’s domain expertise remains indispensable. Although AI can assist with computations, researchers

```
python

from statsmodels.stats.power import TTestIndPower

# Parameters for sample size calculation
effect_size = 0.4 # Estimated effect size
alpha = 0.05 # Alpha level
power = 0.8 # Desired power level
allocation_ratio = 1 # Equal allocation ratio between groups

# Calculate the required sample size per group
analysis = TTestIndPower()
sample_size_per_group = analysis.solve_power(effect_size=effect_size, alpha=alpha,

# Adjust for 10% sample loss
adjusted_sample_size_per_group = sample_size_per_group / 0.9

# Rounding up to nearest whole number since sample size must be an integer
adjusted_sample_size_per_group = int(adjusted_sample_size_per_group)

adjusted_sample_size_per_group

Result
110
```

FIGURE 2 | Code was written by ChatGPT for power analysis based on the given parameters.

must understand the key parameters required for the calculation. For instance, the estimated effect size of 0.4 is to be derived from a literature review or pilot studies. The chosen alpha level of 0.05 and statistical power of 0.8 follow conventional research standards, but the researcher must supply these values. Additionally, the researcher makes the decision to account for a 10% sample loss, and ChatGPT integrates this adjustment into the final calculation.

4.2 | Use Case 6: Descriptive Data Analysis

In the example below, we uploaded simulated data about each orthodontics resident's performance across each key content area covered in our residency curriculum. We then asked ChatGPT to analyse the data to identify areas where the performance is the highest and where it is the lowest.

Prompt: *Generate a performance summary for a non-real group of students in an orthodontic residency program. The summary should include a table listing different disciplines and their corresponding average grades. Additionally, provide an analysis highlighting the best-performing discipline based on average grades, the worst-performing discipline, the best-performing student, and the worst-performing student.*

Output: ChatGPT generated this output by first creating a list of relevant orthodontic residency disciplines and assigning hypothetical numerical grades to each. It then computed the average grade for every discipline and formatted the results into a structured table. To provide further analysis, ChatGPT generated random student names, assigned individual performance

scores, and determined the best and worst performers based on their overall scores. Finally, it summarised the findings in a brief analysis, highlighting the highest and lowest performing disciplines along with the best and worst performing students, ensuring the output was well-organised and easy to interpret, as seen in Figure 3.

4.2.1 | Evaluation

The final evaluation of the process highlights the overall academic performance of orthodontic residency students across multiple disciplines, summarising both individual and subject-based achievements. ChatGPT generated a structured dataset by assigning hypothetical grades to various orthodontic disciplines, calculating the average scores for each subject, and identifying trends in student performance. The results from the figure indicate that Orthognathic Surgery Seminar was the best-performing discipline, with an average score of approximately 3.17, whereas Orthodontic Research Methods I and II followed closely with an average of 2.75. However, Biomechanics in Orthodontics had a slightly lower performance, averaging 2.67, reflecting a moderate challenge for students. Furthermore, an additional analysis determined that Orthodontic Practice Management was the worst-performing discipline, which suggests that students may have struggled more in administrative or business-related aspects of orthodontics compared to clinical and research subjects. On an individual level, Mason emerged as the best-performing student, showcasing consistent academic excellence across subjects, whereas Victoria was identified as the worst-performing student, indicating possible academic difficulties throughout the residency. This evaluation provides valuable insights into

student performance trends, highlighting strengths and areas that may need further curricular improvements.

4.3 | Use Case 7: Data Visualisation

GenAI can also assist with data visualisations. In the following example, we uploaded a spreadsheet containing simulated exam performance data for two cohorts of students—one that received face-to-face instruction before the pandemic (class of 2019) and the other that received remote instruction during the pandemic (class of 2020). We tested and demonstrated ChatGPT's ability to generate a box plot comparing the performance of both cohorts.

Prompt: Create a box plot to visualise students' exam scores that I uploaded. The class of 2019 took the face-to-face version of the course before the pandemic, and the class of 2020 took the online version of the course during the pandemic. I want the class of 2019 to be blue and the class of 2020 to be orange.

Output: Based on the uploaded simulated student exam data (left), ChatGPT illustrated the data, as seen in Figure 4.

4.3.1 | Evaluation

ChatGPT generated the box plot correctly. Taking the Class of 2020 as an example, the horizontal line at the bottom

Performance Summary			
		Discipline	Average Grade
1	Orthognathic Surgery Seminar	Orthognathic Surgery Seminar	3.1666666666666665
2	Orthodontic Research Methods I	Orthodontic Research Methods I	2.75
3	Orthodontic Research Methods II	Orthodontic Research Methods II	2.75
4	Biomechanics in Orthodontics	Biomechanics in Orthodontics	2.6666666666666665

I have provided a detailed performance summary for each discipline. Based on the analysis:

- Best-performing discipline: Orthognathic Surgery Seminar
- Worst-performing discipline: Orthodontic Practice Management
- Best-performing student: Mason
- Worst-performing student: Victoria

FIGURE 3 | Orthodontic residency performance summary (hypothetical data and names).

Class of 2019_F2F	Class of 2020_Online
98	75
74	64
81	93
94	79
73	85
76	84
55	40
83	74
70	32
...	...

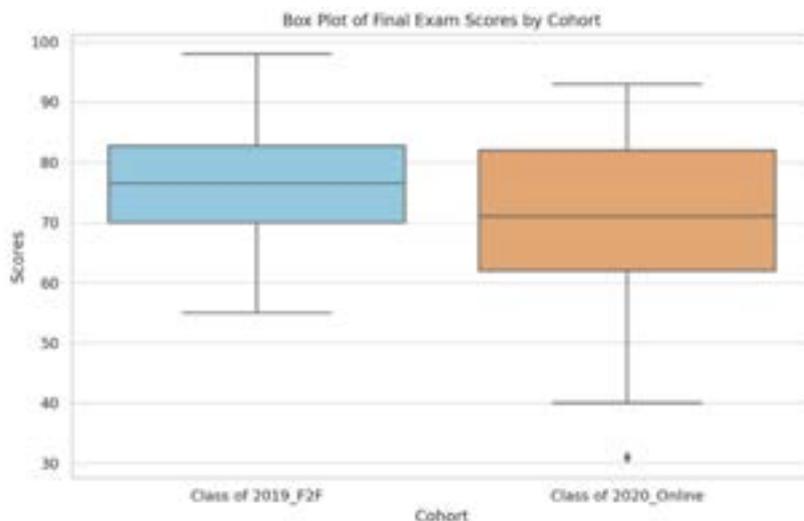


FIGURE 4 | Simulated exam data (left) and ChatGPT generated box plot (right).

represents the minimum score, the line within the box indicates the median, and the line at the top represents the maximum score for the cohort. The blue box encompasses the range where most of the scores fall. Additionally, ChatGPT correctly identified and plotted an outlier score of 32 points as a black diamond. Researchers who are not familiar with the boxplot format can ask ChatGPT to provide an explanation of its components through follow-up prompts. Additionally, researchers can further customise the graphic by using additional prompts to adjust the graphic title, figure size, colours or labels. This example also underscores the importance of researcher-driven decision-making. In this case, we explicitly specified a box plot because it is an appropriate data visualisation based on the research question and dataset. Although AI can assist with data visualisation, it may contain errors in the generated visualisations or produce inappropriate graphic types if the researcher does not specify the desired format or lacks knowledge of which visualisation best fits the research objectives. Therefore, the researcher is responsible for evaluating the usefulness and quality of AI-generated visualisations and ensuring they represent the data in a meaningful way.

4.4 | Use Case Evaluation Metrics

For all the use cases presented, the AI-generated outputs were evaluated qualitatively by the co-authors, who are experienced faculty in dental education. In cases involving specific domains, such as language translation and radiographic interpretation, evaluation was conducted by domain experts. For example, the Mandarin translation was reviewed by a native speaker, and the radiographic content was assessed by a clinician with expertise in dental imaging. Although this manuscript does not apply formal benchmarking tools or scoring rubrics, the outputs were assessed based on clarity, educational value, contextual relevance, and clinical plausibility. This is a limitation of this study; however, it is within the scope of this narrative review.

5 | Discussion

As demonstrated in several examples from earlier sections, data privacy and confidentiality are a critical challenge when using AI tools for data-related projects. Researchers and educators must comply with ethical and legal standards, such as the Health Insurance Portability and Accountability Act (HIPAA) and Family Educational Rights and Privacy Act (FERPA), when handling healthcare and educational data. In the examples above, we used simulated (fake) data for demonstration purposes to ensure privacy protection. However, in real-life applications, a recommended solution is to ask the AI tool to generate a script or code that can be downloaded and run locally in a secure environment. Instead of uploading data to an AI platform, researchers with some programming knowledge can modify their prompt to include the request ‘Write the code that I can download and run locally.’ By executing the AI-generated script in a local coding environment such as Visual Studio Code (VS Code), researchers can mitigate the risk of data exposure while ensuring compliance with ethical and legal requirements.

By adopting these local execution strategies, researchers can enhance data protection while still leveraging the capabilities of generative AI. However, as the use of genAI tools expands beyond research workflows into dental education and academic practice, additional challenges arise that demand careful consideration and mitigation, such as: (a) Data Privacy and Security: The use of genAI tools within healthcare and educational contexts introduces concerns around the handling of protected health information and student data, particularly when using platforms that process information on external servers. This creates potential conflicts with regulations such as HIPAA and FERPA, as discussed in the previous section. As noted by Jin et al. [22] institutional policies have not yet caught up with the pace of AI adoption, leaving gaps in data governance and compliance. (b) Misinformation and Hallucinations: Large language models are known to produce inaccurate or fabricated content, often presented with high confidence. This so-called ‘hallucination’ effect poses risks in learning environments, where students may accept flawed outputs as fact. Kung et al. [9] demonstrated that ChatGPT occasionally produced erroneous answers during medical licensing simulations, emphasising the need for critical AI literacy and content verification mechanisms. (c) Diagnostic Accuracy and Clinical Judgement: While genAI holds the potential in augmenting diagnostic support, it lacks clinical reasoning and contextual sensitivity. Over-reliance on AI-generated interpretations—particularly in image-based tasks—may lead to incorrect treatment decisions. Ali et al. [13] found that ChatGPT underperformed in tasks requiring nuanced image evaluation in dental education. These limitations highlight the importance of maintaining strong human oversight and ensuring genAI tools are used as supportive aids rather than decision-making authorities.

GenAI tools can assist with a variety of tasks, as demonstrated in the use cases in this paper. However, these examples also highlight their limitations, such as inaccurate interpretation of exam reports. Given these limitations, human oversight is critical to ensure they are used responsibly and safely. These tools should be used as a first step in the process rather than as the final authority. Our use cases also highlight that the benefits of AI rely on human expertise and creativity. Its potential is maximised when users have domain knowledge to effectively guide AI with relevant prompts, evaluate response quality, and make informed decisions. Finally, whereas faculty and staff can use AI tools to assist with daily tasks, we must also guide students to use AI as a learning aid rather than a shortcut, as demonstrated in the examples. This ensures they develop the skills needed to navigate a future where AI will play a significant role in both their personal and professional lives.

6 | Conclusion

In conclusion, AI integration must be approached strategically at both the individual and institutional levels. Institutions should provide targeted training to help faculty, staff and students navigate AI’s capabilities, limitations and responsible use. Additionally, clear guidelines and policies must be developed to ensure ethical implementation, particularly concerning data privacy, confidentiality, academic integrity and legal compliance. As AI tools become more embedded in education, institutional

readiness and faculty leadership will be key to ensuring their safe and effective integration.

Author Contributions

All authors contributed to the development of this review. J.B. and M.Z. were responsible for conducting the literature search, selecting relevant studies and drafting the initial manuscript. J.B. and M.Z. provided feedback and assisted in refining the final manuscript. All authors have reviewed and approved the final version.

Acknowledgements

This study was partially funded by Orthodontic Faculty Development Fellowship Awards (OFDFA)—Robert E. Binder Teaching Fellowship Award.

Disclosure

The authors have nothing to report.

Ethics Statement

The authors have nothing to report.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The authors have nothing to report.

References

1. Google, "What Is Artificial Intelligence (AI)? *Google Cloud*," (2024), <https://cloud.google.com/learn/what-is-artificial-intelligence>.
2. Company M&, "What Is Generative AI? *McKinsey Explainers*," (2024), <https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-generative-ai>.
3. K. Singhal, S. Azizi, T. Tu, et al., "Large Language Models Encode Clinical Knowledge," *Nature* 620, no. 7972 (2023): 172–180.
4. H. Wang, A. Dang, Z. Wu, and S. Mac, "Generative AI in Higher Education: Seeing ChatGPT Through Universities' Policies, Resources, and Guidelines," *Computers and Education: Artificial Intelligence* 7 (2024): 100326.
5. B. L. Moorhouse, M. A. Yeo, and Y. Wan, "Generative AI Tools and Assessment: Guidelines of the World's Top-Ranking Universities," *Computers and Education Open* 5 (2023): 100151.
6. A. Darvishi, H. Khosravi, S. Sadiq, D. Gašević, and G. Siemens, "Impact of AI Assistance on Student Agency," *Computers in Education* 210 (2024): 104967.
7. J. Cordero, J. Torres-Zambrano, and A. Cordero-Castillo, "Integration of Generative Artificial Intelligence in Higher Education: Best Practices," *Education in Science* 15, no. 1 (2024): 32.
8. L. Mohebi, "Empowering Learners With ChatGPT: Insights From a Systematic Literature Exploration," *Discover Education* 3, no. 1 (2024): 36.
9. T. H. Kung, M. Cheatham, A. Medenilla, et al., "Performance of ChatGPT on USMLE: Potential for AI-Assisted Medical Education Using Large Language Models," *PLOS Digital Health* 2, no. 2 (2023): e0000198.
10. C. K. Lo, "What Is the Impact of ChatGPT on Education? A Rapid Review of the Literature," *Education in Science* 13, no. 4 (2023): 410.
11. C. Preiksaitis and C. Rose, "Opportunities, Challenges, and Future Directions of Generative Artificial Intelligence in Medical Education: Scoping Review," *JMIR Medical Education* 9 (2023): e48785.
12. P. Furey, A. Town, K. Sumera, and C. A. Webster, "Approaches for Integrating Generative Artificial Intelligence in Emergency Healthcare Education Within Higher Education: A Scoping Review," *Critical Care Innovations* 7 (2024): 34–54.
13. K. Ali, N. Barhom, F. Tamimi, and M. Duggal, "ChatGPT—A Double-Edged Sword for Healthcare Education? Implications for Assessments of Dental Students," *European Journal of Dental Education* 28, no. 1 (2024): 206–211.
14. C. Chakraborty, S. Pal, M. Bhattacharya, S. Dash, and S. S. Lee, "Overview of Chatbots With Special Emphasis on Artificial Intelligence-Enabled ChatGPT in Medical Science," *Frontiers in Artificial Intelligence* 6 (2023): 1237704.
15. L. Laranjo, A. G. Dunn, H. L. Tong, et al., "Conversational Agents in Healthcare: A Systematic Review," *Journal of the American Medical Informatics Association* 25, no. 9 (2018): 1248–1258.
16. A. S. Miner, A. Milstein, S. Schueller, R. Hegde, C. Mangurian, and E. Linos, "Smartphone-Based Conversational Agents and Responses to Questions About Mental Health, Interpersonal Violence, and Physical Health," *JAMA Internal Medicine* 176, no. 5 (2016): 619–625.
17. J. E. Bibault, B. Chaix, P. Nectoux, A. Pienkowski, A. Guillemasé, and B. Brouard, "Healthcare ex Machina: Are Conversational Agents Ready for Prime Time in Oncology?," *Clinical and Translational Radiation Oncology* 16 (2019): 55–59.
18. J. H. Gruenhagen, P. M. Sinclair, J. A. Carroll, P. R. A. Baker, A. Wilson, and D. Demant, "The Rapid Rise of Generative AI and Its Implications for Academic Integrity: Students' Perceptions and Use of Chatbots for Assistance With Assessments," *Computers and Education: Artificial Intelligence* 7 (2024): 100273.
19. D. O. Eke, "ChatGPT and the Rise of Generative AI: Threat to Academic Integrity?," *Journal of Responsible Technology* 13 (2023): 100060.
20. M. McCormack, "EDUCAUSE QuickPoll Results: Adopting and Adapting to Generative AI in Higher Ed Tech. EDUCAUSE Research Notes," 2023.
21. F. Miao and W. Holmes, *Guidance for Generative AI in Education and Research* (UNESCO, 2024), <https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research>.
22. Y. Jin, L. Yan, V. Echeverria, D. Gašević, and R. Martinez-Maldonado, "Generative AI in Higher Education: A Global Perspective of Institutional Adoption Policies and Guidelines," *Computers and Education: Artificial Intelligence* 8 (2025): 100348.
23. K. Lucariello, *Survey Suggests Higher Ed Institutions Are Not Ready for Generative AI* (Campus Technology, 2024), https://campustechnology.com/Articles/2024/02/22/Survey-Suggests-Higher-Ed-Institutions-Are-Not-Ready-for-Generative-AI.aspx?s=ct_ctow_190324&oly_enc_id=3970E5156245B8W.