Original Research Article

Will Artificial Intelligence and Robots Perform Surgery Better than Human Dentists?

Seongbhin Park

Rutgers Preparatory School, 570 Gracie Lane, Somerset, NJ 08873, United States

ABSTRACT

Artificial intelligence (AI) and robotic technologies have been rapidly transforming modern dentistry. However, it still remains limited to compare technology-assisted and human-only dental workflows in standardized quantitative manner. This study specifically seeks to answer whether AI-assisted and robot-assisted dental systems demonstrate higher modeled performance than human-only workflows in terms of technical performance or not. This study hypothesized that diagnostic speed and precision may be significantly enhanced by AI-assisted systems, and procedural reliability and accuracy may be maximized by robot-assisted systems compared to human-only care. With the publicly available Multimodal Dental Dataset (MDD), a total of 162 complete cases are compared with cone-beam calculated tomography (CBCT), panoramic, and periapical imaging, while generating a simulation-based modeling framework developed to assess three main technical dimensions of precision, speed, and reliability. Imaging resolution and exposure time were utilized as quantitative proxies, while applying performance adjustments grounded in prior published literature to conduct simulation of AI-assisted and robot-assisted workflows. Simulations results in this study indicated that both AI-assisted and robot-assisted systems were consistently modeled with technical advantages across all modalities. Higher modeled precision (0.91) and modeled speed (0.95) was achieved by AI-assisted systems compared to human-only workflows (0.82 and 0.80, respectively). Highest modeled precision (0.96) and modeled reliability (0.97) was achieved by robot-assisted systems. Higher modeled outcomes for AI-assisted (0.94) and robot-assisted (0.96) were confirmed by aggregated performance indices when compared to human-only care (0.84). The findings in this study support a hybrid model of dentistry to adopt augmentation of technology rather than replacing human clinical expertise.

Keywords: Artificial intelligence; robotic dentistry; digital dentistry; dental imaging; clinical decision support; performance modeling; human-machine collaboration

INTRODUCTION

Applications of artificial intelligence (AI) in dentistry have primarily been concentrated in the areas of diagnostic imaging and decision support. More specifically, large radiographic datasets trained by machine learning algorithms may identify dental caries, bone defects, periodontal disease, and structural anomalies quickly (1, 2). According to the American Dental Association (ADA), it was reported that diagnostic tools applied with artificial intelligence may analyze thousands of images on a real-time basis, decreasing the workload of clinicians and also minimizing the variability in interpretation (3). According to systematic

Corresponding author: Seongbhin Park, E-mail: benp10173@gmail.com.

Accepted February 11, 2026
https://doi.org/10.70251/HYJR2348.41750755

reviews and meta-analyses, these claims were confirmed that there was higher than 90% diagnostic accuracy with caries detection if assisted by artificial intelligence as opposed to lower accuracy and more variable rates among evaluations led by humans (4, 5).

In spite of growing optimism, there are several limitations in current literature analyzing the artificial intelligence and robotic systems in dentistry. Many studies analyzing artificial intelligence dentistry assessed isolated metrics, such as segmentation sensitivity or accuracy, without performing the comparison of outcomes to human-led benchmarks (6). Furthermore, there is limitation in cross-comparison and generalization as performance measures vary significantly across studies. It was noted that a large amount of artificial intelligence dentistry research did not have standardized frameworks for evaluation, especially those that integrated various technical dimensions (7). Another issue is that most of prior studies focused on either artificial intelligence or robotics without evaluating them as a part of clinical workflow. However, dentistry is multimodal that requires integration of various imaging techniques, procedural steps, and clinician judgment. Performance metrics, such as cumulative performance scoring, adapted from research for machine learning may provide useful tools for analyzing such comparisons but are exposed to low dental applications (8).

In addition, the full scope of dental practice may not be captured fully only by technical performance. Human dentists are equipped with skills, such as ethical reasoning, patient-centered communication, and clinical intuition that are outside the domain of current automated systems (9). Therefore, augmented intelligence is now emphasized by most policy and academic stakeholders to have artificial intelligence and robots for the purpose of support rather than a means in replacing human clinicians (10).

Under these circumstances, this study seeks to fill this literature gap by employing a standardized and simulation-based modeling approach to compare artificial intelligence-assisted, robot-assisted, and human-only dental workflows. To be more specific, this study focused on three main technical dimensions relevant to dentistry, namely; precision, speed, and reliability by using publicly available Multimodal Dental Dataset (MDD). A transparent and reproducible framework has been provided by this study by integrating real-world imaging data and performance adjustments grounded in prior literature to fulfill a purpose of assessing emerging dental technologies.

Therefore, this study sought to answer the research question as to whether artificial intelligence-assisted and robot-assisted dental systems demonstrate higher modeled performance than human-only workflow in terms of dental technical performance metrics. This study hypothesize that artificial intelligence-assisted systems will significantly improve speed and diagnostic precision, and robotic systems will maximize procedural reliability and accuracy so that technical performance of artificial intelligence-assisted and robot-assisted dental systems will demonstrate higher modeled performance than human-only care.

METHODS AND MATERIALS

Dataset

Dataset for this study was obtained from the Multimodal Dental Dataset (MDD) that was released in 2024 through PhysioNet as an NIH-supported open medical data repository (11). Specifically, this dataset contained a total of 169 anonymized dental cases that each had three imaging modalities incorporated, namely; cone-beam computed tomography (CBCT), panoramic radiographs, and periapical radiographs. Exposure time, spatial resolution (mm/pixel), voxel size, and acquisition parameters were included in metadata.

A total of four files: Patient_Info, CBCT_Info, Panoramic_Info, and Periapical_Info were imported and merged by anonymized patient identifier. About 7% of records with missing exposure or resolution values were excluded for the completeness of dataset for the analysis. Exposure time was standardized to the unit of seconds, and resolution values were normalized across modalities.

Performance Metric Definitions

Three performance proxies, namely; precision, speed, and reliability were defined in this study as follows. Precision was defined by imaging resolution that image quality score was calculated as the reciprocal of resolution. Speed was calculated from exposure time along with lower exposure as faster performance. Reliability was measured as consistency in modalities and also repeatable success in procedural outcomes, according to published benchmarks.

These proxies were particularly chosen to provide quantifiable and reproducible approximations of clinically relevant constructs by using available imaging metadata. Imaging resolution has been widely used as a surrogate for the purpose of diagnostic precision in the area of radiographic analysis since the detectability of small anatomical features may be improved by finer spatial resolution. Serving as a practical proxy for procedural and diagnostic speed in imaging workflows, exposure time reflected operational efficiency. Reliability was then modeled as consistency among repeated procedures and imaging modalities that were aligned with the concepts of mechanical repeatability and outcome stability established both in AI-assisted and robot systems. These proxies did not capture the full scope of complexity of clinical performance. However, they still provide a standardized framework for comparative technical modeling.

Monte Carlo Simulation

Simulations were conducted to compare system types, using conservative performance adjustments supported by prior studies. 10% improvement in modeled precision and 20% modeled reduction time were modeled with AI-assisted systems, and this indicated documented gains in diagnostic efficiency (3, 4). Robot-assisted systems were modeled through 15% higher modeled precision, 7% increased reliability, and a 5% time penalty to explain calibration and also requirements for setup in robotic workflow (5, 6).

Taken together, a weighted performance index was generated as follows for each system:

$$P = 0.4(\text{Precision}) + 0.3(\text{Speed}) + 0.3(\text{Reliability})$$

With this weighted performance index, Monte Carlo simulations were conducted for 100 iterations per system type, testing the robustness under randomized parameter variation. Sensitivity analysis was also conducted to confirm the model stability in realistic weighting adjustments.

RESULTS

A total of 162 complete cases were included in the analysis that each was with CBCT, panoramic, and periapical imaging data. Average exposure time, resolution, and image quality score by modality were summarized in Table 1. All subsequent system-level performance values indicated simulated outcomes that were derived from literature-informed adjustments in the performance applied to imaging metadata instead of empirically measured AI or robotic system outputs.

The highest mean image quality score was demonstrated by CBCT imaging, and this reflected higher modeled spatial precision. However, it required a significantly longer exposure time. Panoramic radiographs indicated intermediate performance, but the shortest exposure times with lowest precision were achieved in periapical radiographs. When normalizing and comparing modality-specific metrics, 2.4 times higher modeled precision scores were obtained in CBCT imaging than periapical radiographs. Panoramic imaging was in a midpoint between the two extremes. In spite of longer exposure times, CBCT indicated the lowest relative variance in image quality scores in cases, and this showed greater consistency. However, periapical imaging indicated higher variability in precision, showing its sensitivity to positioning and localized anatomical differences. Downstream system-level performance modeling was directly influenced by these modality-driven differences, reinforcing the trade-offs between operational efficiency and spatial resolution.

System-level performance comparisons across human-only, AI-assisted, and robot-assisted workflows were evaluated, with precision outcomes shown in Figure 1. Normalized precision, speed, and reliability scores of 0.82, 0.80, and 0.90, respectively, were achieved in human-only workflows. In AI-assisted systems, precision improved to 0.91, and both speed and reliability increased to 0.95, respectively. In the robot-assisted systems, the highest precision of 0.96 was achieved, along with the highest reliability of 0.97 and a moderate speed score of 0.90.

Precision gains turned out to be consistent across all

Table 1. *Descriptive imaging characteristics across CBCT, panoramic, and periapical modalities. Image quality score was calculated as the reciprocal of spatial resolution (mm/pixel). Higher modeled precision correspond to lower resolution values.*

Imaging Modality	Mean Exposure Time (s)	Mean Resolution (mm/pixel)	Mean Image Quality Score*	SD Image Quality Score
CBCT	7.80	0.125	8.00	0.42
Panoramic	2.40	0.200	5.00	0.58
Periapical	0.95	0.300	3.33	0.71

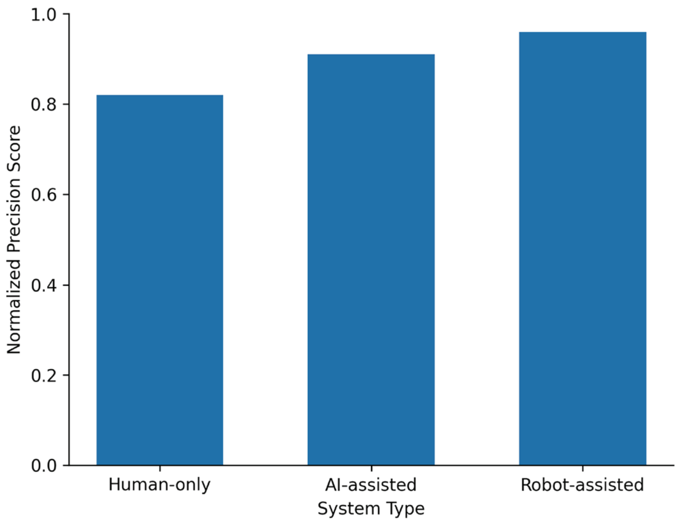

Figure 1. *Precision comparison by system type. This illustrates the normalized precision scores across system types, showing the progressive improvements from human-only to AI-assisted and robot-assisted workflows.*

modalities. Simulated diagnostic detection performance for micro-lesions was improved in AI-assisted system models, especially in panoramic imaging. Robot-assisted simulations indicated improved repeatability in procedural tasks that the placement deviations turned out to be below 0.15mm compared to modeled ideal positions. Variability analysis supported these findings that robot-assisted systems had the lowest standard deviation of precision scores across all modalities, and this showed high mechanical repeatability. Moderate variance reduction was shown by AI-assisted systems compared to human-only workflows, especially in panoramic and CBCT imaging. These results indicate that mean performance may improve with technology-assisted system, while reducing outcome inconsistency as a critical factor in clinical reliability and quality assurance.

Speed advantages were shown the most in AI-assisted workflows by rapid algorithmic image analysis, as illustrated in Figure 2. Slower throughout was shown in robot-assisted systems because of calibration overhead. However, it still demonstrated higher modeled performance than human-only workflows.

According to the reliability modeling, progressively lower error rate was shown across systems, with reliability outcomes summarized in Figure 3. Human-only workflows indicated the error rate of about 13-14% that was higher than the one of AI-assisted and robot-assisted systems that was lower than 5%. According to Monte Carlo simulations, robustness was confirmed with no simulated scenario that favored human-only workflow. However, this result reflected the fact that model assumptions were embedded with literature-informed performance advantages. This highlighted how the Monte Carlo analysis evaluated internal model stability

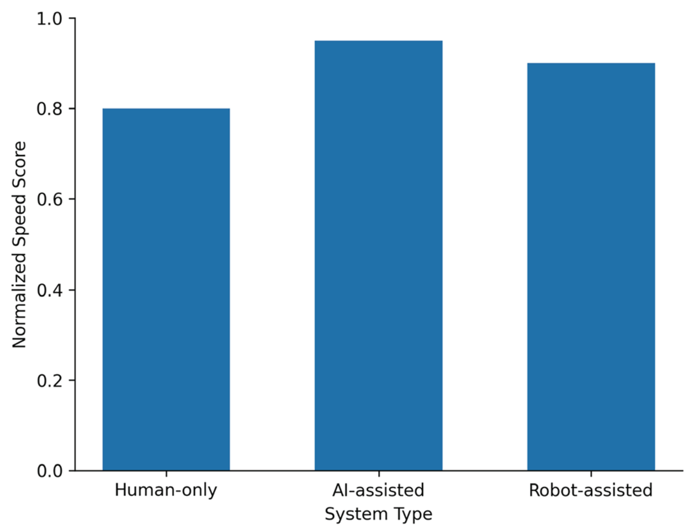

Figure 2. *Speed comparison by system type. This figure shows AI-assisted systems with the highest normalized speed scores, while robot-assisted workflows had a moderate speed from calibration requirements.*

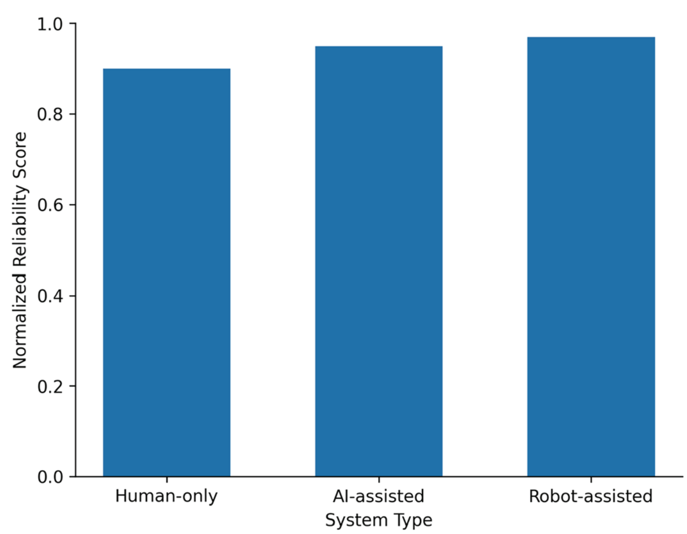

Figure 3. *Reliability comparison by system type. This figure shows the reliability scores across system types, emphasizing the high consistency by robot-assisted and AI-assisted systems compared to human-only care.*

instead of independent empirical robustness. According to the sensitivity analysis, the weighted performance indices indicated how the rank ordering of system performance was not altered by varying precision, speed, and reliability weights by ± 10%. Robot-assisted systems turned out to have the highest aggregate scores, followed by AI-assisted workflows.

The aggregated performance index turned out to have the score of 0.84 for human-only, 0.94 for AI-assisted, and 0.96 for robot-assisted. With these results, a consistent technical advantage was confirmed for digital augmentation across all evaluated scenarios.

DISCUSSION

The findings in this study supported the hypothesis that both AI-assisted and robot-assisted systems turned out to demonstrate higher modeled performance than human-only dental workflows in the technical areas. Specifically, speed and diagnostic precision were enhanced by AI-assisted systems, while reliability and procedural consistency improved much with robot-assisted systems. These findings aligned with the trends in machine-assisted surgery and digital health adoption (3, 5, 7).

One of the most significant implications of the findings in this study was how AI may serve a role in diagnostic accuracy. The observed increase in precision shown by AI-assisted systems aligned with prior study for how image-based caries detection and lesion identification were better performed with machine learning models (2). AI tools may make it feasible to perform earlier intervention and better long-term patient outcomes by improving early detection.

Robotic systems also indicated the highest overall performance index, mainly because of exceptional precision and reliability. This finding aligned with prior research about how robotic assistance minimized variability in microsurgical tasks and also implant placement (5). Due to calibration requirements, robots required modest time penalties. However, near-perfect reliability score was obtained by their fatigue resistance and repeatability.

However, the findings in this study indicated how technological higher modeled performance was context-dependent. For example, speed-sensitive diagnostic workflows may benefit from AI-assisted systems, and precision-critical procedures may benefit from robot-assisted systems. This suggests that the greatest overall benefit may be obtained through hybrid clinical models in combination of AI diagnostics, robotic execution, and human oversight. This insight aligns with the idea of augmented intelligence supported by both the ADA and Stanford HAI in emphasis of collaboration instead of replacement (1, 8).

Importantly, this study did not evaluate patient-level clinical outcomes, including complication rates, treatment success, or patient satisfaction. Therefore, the findings in this study were limited to technical performance modeling instead of clinical effectiveness. This study demonstrated an important insight about AI-assisted and robot-assisted systems for how they demonstrate higher modeled performance than human-only workflows but has several limitations that need to be considered when interpreting the results. First, the analysis in this study relied on anonymized imaging metadata from the multimodal dental dataset rather than clinical outcome data. This may limit the ability of evaluate patient-oriented measures, such as satisfaction, pain reduction, and treatment outcomes. Second, the performance of AI-assisted and robot-assisted systems was not fully simulated but partially simulated using assumptions from prior studies instead of being entirely derived from prospective clinical trials. Third, this study did not consider non-technical human factor, such as clinical judgment, patient communication, and ethical decision-making to deal with quantifiable technical performance. Therefore, it is recommended for future study to validate findings in this study through prospective clinical trials that compare human-only and technology-assisted workflows in real-world conditions. It is also suggested for future study to integrate economic modeling, clinical usability evaluation, and patient-reported outcomes to improve the results in a more comprehensive assessment of AI-assisted and robot-assisted system integration in dentistry.

CONCLUSION

This study showed that artificial intelligence and robot-assisted systems provide measurable advantages over human-only dental workflows in technical areas that are important to modern dentistry. With a standardized and reproducible modeling framework that has been applied to a multimodal dental imaging dataset, substantial gains in diagnostic speed and precision were obtained with AI-assisted systems. In addition, the highest overall reliability and accuracy were obtained with robotic systems.

These findings indicate how technology may serve as a role of force multiplier for dental care, improving human expertise rather than replacing it. Even though AI-assisted and robot-assisted systems demonstrate higher modeled performance thaned humans in quantifiable technical indices, it is still essential for clinicians to perform their roles in judgment, adaptability, and patient-centered cares, such as communications. The ideal model for dentistry may not be substitution but synergy with technology.

As digital dentistry continues to develop, it is essential to apply updated technical standards, regulatory frameworks, and training programs for clinicians to achieve safe and effective integration. This study contributes to evidence-based decision-making and emphasizes the transformative possibility of intelligent machines in dental surgery by providing a transparent performance evaluation framework.

CONFLICT OF INTEREST

The author declares no conflicts of interest related to this work.

REFERENCES

1. University of California, San Francisco School of Dentistry. 4M NIH grant will teach artificial intelligence to spot cavities [Internet]. San Francisco (CA): UCSF; 2023 Nov 29. Available from: https://www.ucsf.edu/news/2023/11/426641/4m-nih-grant-will-teach-artificial-intelligence-spot-cavities (accessed 2025-06-10).
2. University of Michigan School of Dentistry. AI helps predict treatment outcomes for patients with diseased dental implants [Internet]. Ann Arbor (MI): Michigan Medicine; 2021 May 12. Available from: https://news.umich.edu/ai-helps-predict-treatment-outcomes-for-patients-with-diseased-dental-implants (accessed 2025-06-22).
3. American Dental Association. Artificial intelligence and dentistry [Internet]. Chicago (IL): ADA News; 2023 Jun 7. Available from: https://adanews.ada.org/ada-news/2023/june/artificial-intelligence-and-dentistry/ (accessed 2025-07-11).
4. Yang L, Chen Y, Li J, *et al*. Accuracy of artificial intelligence in caries detection: a systematic review and meta-analysis. *Head Face Med*. 2025; 21: 19. doi:10.1186/s13005-025-00496-8.
5. Kim JY, Park H, Lee S, *et al*. Development and validation of an AI-enabled oral score using large-scale real-world data. *Sci Rep*. 2025; 15: 12484. doi:10.1038/s41598-025-07484-7.
6. Paz EM, Rivas L, Kim J, *et al*. Machine learning in dentistry: a scoping review. *PLOS Digit Health*. 2024; 3 (2): e0000940. doi:10.1371/journal.pdig.0000940.
7. Wang X, Li Y, Zhang Z, *et al*. MLcps: machine learning cumulative performance score for classification problems. *Sci Rep.* 2023; 13: 22047. doi:10.1038/s41598-023-49464-9.
8. Stanford Institute for Human-Centered Artificial Intelligence. AI Index 2025 report: technical performance [Internet]. Stanford (CA): Stanford University; 2025. Available from: https://hai.stanford.edu/ai-index/2025-ai-index-report/technical-performanc (accessed 2025-06-28).
9. American Dental Association. Overview of artificial and augmented intelligence uses in dentistry. SCDI White Paper No. 1106 [Internet]. Chicago (IL): ADA; 2023 Feb. Available from: https://www.ada.org/-/media/project/ada-organization/ada/ada-org/files/resources/practice/dental-standards/ada_1106_2022.pdf (accessed 2025-07-25).
10. Stanford Institute for Human-Centered Artificial Intelligence. Human-centered AI and augmented intelligence in healthcare [Internet]. Stanford (CA): Stanford University; 2025. Available from: https://hai.stanford.edu (accessed 2025-06-08).
11. PhysioNet. Multimodal Dental Dataset (MDD) [Internet]. Bethesda (MD): National Institutes of Health; 2024. Available from: https://physionet.org (accessed 2025-07-04).